



# Durham E-Theses

---

## *Interpersonal deceit and lie-detection using computer-mediated communication*

Placks, Simon James

### How to cite:

---

Placks, Simon James (2003) *Interpersonal deceit and lie-detection using computer-mediated communication*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/3707/>

### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Interpersonal Deceit and Lie-Detection Using  
Computer-Mediated Communication

Simon James Placks

Ph.D. Thesis

2003

Department of Computer Science  
University of Durham

*A copyright of this thesis rests  
with the author. No quotation  
from it should be published  
without his prior written consent  
and information derived from it  
should be acknowledged.*



19 JAN 2004

---

# Abstract

This thesis examines the use of computer-mediated communication for lie-detection and interpersonal deceit. The literature within the fields of lie-detection and mediated communication are reviewed and it is proposed that there is a lack of knowledge surrounding how people use CMC to deceive one another. Qualitative research was carried out in order to address this shortcoming, exploring the self-reported experiences of chat room users who have been exposed to online deceit. Reports were provided that describe the misrepresentation of age, gender, vocation, affection, and appearance. The importance of stereotypes in driving suspicions is also emphasised within the reports. It is suggested that this key characteristic has more dominance in CMC than it would do face-to-face because of the occlusion of the traditional nonstrategic clues to deceit. Evidence for an alternative set of nonstrategic leakage clues was examined further by conducting a variant of the Guilty-Knowledge test within the context of a CMC based crime. It was found that participants exhibited a response time inhibition effect when presented with 'guilty knowledge' and that this effect was detectable through a standard two-button mouse. The use of such nonstrategic cues to deceit was explored further in a study that examined how CMC might be used to add additional control to a Statement Validity Assessment truth-validation test. It was found that the content analysis technique used by SVA was unable in its present form to correctly distinguish between truthful and fabricated statements of participants interviewed using a CMC chat program. In addition, it was found that the deletion-behaviours of participants fabricating a story within CMC provided no quantitative or qualitative evidence that they were lying.

---

## Copyright Notice

The copyright of this thesis rests with the author. No quotation from it should be published without their prior written consent and information derived from it should be acknowledged.

## Declaration

No part of this material has previously been submitted for a higher degree at the University of Durham or at any other University.



---

# Acknowledgements

None of this work would have been possible without the copious support of numerous friends, relatives, and colleagues. I would like to thank my parents Vivien and Michael Placks for their constant love and support, and my grandparents Reg, Freda, and June. A special thank-you also goes to my supervisor, Dave Robson, for his help and guidance during the development of this research.

I would like to thank all the staff and postgraduates of the Department of Computer Science for being fantastic in every way. Another special thank-you goes to Lucy Akehurst for her help and support. I would also like to thank Rachel Mather, Paul Johnson, Nick Davis, Iain Sinclair, Ian Graham, Elliot Lawler, Katherine Tyas, Stephan Jamieson, Andy Hatch, Mike Smith, Chris Taylor, Claire Knight, and Robert Kiesling for generally being a great bunch of people. Thanks also go to Tim Burman for helping me construct and equip my laboratory. Much of this work would not have been possible without a great number of unnamed participants, raters and respondents, all to whom I am indebted. I would also like to thank the Durham University Ethics Committee for their helpful comments and suggestions. Finally, I would like to thank my examiners Malcolm Munro and Gillian Lovegrove for the time spent reviewing this thesis and for their constructive critiques and comments.

---

# Table of Contents

## **Chapter 1: Introduction**

1.1	The Problem.....	1
1.2	Computer-Mediated Communication.....	2
1.3	Lie Detection.....	2
1.4	Research Questions.....	3
1.5	Thesis Overview.....	4
1.6	Notes.....	5

## **Chapter 2: Interpersonal Deception Theory and the Semantics Of Deceit**

2.1	Searching for a Definition of a Lie.....	7
2.2	Categorising Lies.....	8
2.2.1	Criteria of Categorisation.....	8
2.2.1.1	Lie Complexity.....	8
2.2.1.2	Outcome.....	9
2.2.1.3	Moral Turpitude.....	9
2.2.1.4	Cultural Factors.....	10
2.2.1.5	Lie Mechanics.....	10
2.2.1.6	Actors.....	10
2.2.1.7	Motivation.....	11
2.2.1	Categorising the Lies of Children.....	11
2.3	Interpersonal Deception Theory.....	12
2.3.1	The Theory.....	12
2.3.1.1	The Propositions of IDT.....	14
2.3.2	Critique.....	16
2.3.3	Empirical Support for IDT.....	18
2.4	Summary.....	19

## **Chapter 3: Detecting Deceit Through the Analysis of Verbal and Physiological Cues**

3.1	Statement Validity Assessment.....	22
3.1.1	History of SVA.....	22
3.1.2	Overview of Methodology.....	25
3.1.3	Overview of Empirical Studies.....	30
3.1.3.1	The Experimental Approach.....	30

3.1.3.2 The Field Study Approach.....	32
3.1.4 Interviewing for SVA: Research Issues.....	33
3.1.4.1 Principles Underlying the SVA Interview.....	33
3.1.4.2 The Cognitive Interview.....	35
3.1.5 CBCA and the Validation Checklist: Research Issues.....	37
3.1.5.1 Theoretical Support.....	37
3.1.5.2 Training.....	37
3.1.5.3 Establishing Ground Truth.....	39
3.1.5.4 The Base Rate of Lying.....	40
3.1.5.5 Reliability.....	40
3.1.5.6 Realism and Construct Validity.....	42
3.1.5.7 External Validity.....	43
3.1.5.8 Criteria Weighting and Configuration.....	44
3.1.5.9 New Criteria.....	44
3.2 Reality Monitoring.....	45
3.2.1 Background.....	45
3.2.2 The Criteria.....	46
3.2.3 Empirical Evidence.....	46
3.3 Other Verbal Approaches.....	47
3.3.1 SCAN.....	47
3.3.2 Lexical Diversity.....	48
3.4 The Polygraph.....	49
3.4.1 The Control Questions Test.....	49
3.4.2 The Guilty Knowledge Test.....	52
3.5 Summary.....	53

#### **Chapter 4: Computer-Mediated Communication**

4.1 The Origins of CMC Technology and Research.....	55
4.2 Theoretical Approaches to CMC Research.....	57
4.2.1 The Social Presence Approach.....	57
4.2.2 The Filtered-Cues Paradigm.....	58
4.2.3 The SIDE Model.....	60
4.2.4 The Social Information Processing Paradigm.....	63
4.2.5 An Ecological Approach.....	66
4.3 The Structural Features of Quasi-Synchronous CMC.....	68
4.4 The Self Online.....	70
4.5 Deception in CMC.....	73
4.5.1 Case Studies.....	73
4.5.2 Recent Research.....	75

---

4.6 Summary.....	76
 <b><u>Chapter 5: Exploring Interpersonal Deceit, Lie-Detection, and CMC</u></b>	
5.1 Lie-Detection and CMC.....	78
5.2 A Process for a Practitioner.....	80
5.2.1 Regulation and Investigation.....	80
5.2.2 Examination of the Social Context.....	80
5.2.3 Examination of the Technological Context.....	81
5.2.4 The Prevalence of Innocuous or Detrimental Deceit.....	81
5.2.5 Selecting a Detection Strategy.....	81
5.2.6 Present Difficulties .....	82
5.3 The Interpersonal Deceptive Process in CMC.....	82
5.3.1 Context Interactivity and Social Cues.....	84
5.3.2 Perceptual Asynchrony.....	85
5.3.3 Multiplicity of Interaction.....	85
5.3.3.1 Salient Group Identity.....	86
5.3.3.2 Salient Individual Identity.....	86
5.3.4 The Reconstruction of Speaker and Listener Roles.....	86
5.3.5 The Lack of Research into Non-Strategic Leakage.....	87
5.4 The Non-Strategic Deceptive Behaviours of CMC.....	87
5.4.1 Cognitive Load.....	88
5.4.2 Affect: Duping Delight.....	89
5.4.3 Affect: Detection Apprehension.....	90
5.4.4 Arousal.....	90
5.4.5 Attempted Control.....	90
5.4.6 A Caution.....	91
5.5 The Strategic Deceptive Behaviours of CMC.....	91
5.5.1 Method.....	92
5.5.2 Analysis.....	94
5.5.2.1 Gender.....	94
5.5.2.2 Age.....	96
5.5.2.3 Relationship Status and Affection.....	97
5.5.2.4 Vocation and Skills.....	98
5.5.2.5 Appearance.....	99
5.5.3 Evaluation.....	99
5.6 A Semantic Scheme for QS-CMC Lie-Detection.....	100
5.7 Summary and Discussion.....	101

---

## **Chapter 6: Response Time Measures to Assess ‘Guilty Knowledge’ within a CMC Scenario**

6.1 Background.....	103
6.2 Hypotheses.....	106
6.3 Method.....	106
6.3.1 Participants.....	107
6.3.2 Design and Protocol.....	107
6.3.2.1 Arrival.....	107
6.3.2.2 Briefing.....	107
6.3.2.3 Mock Crime.....	108
6.3.2.4 Distracter Task.....	109
6.3.2.5 Testing Phase.....	109
6.3.2.6 Debrief.....	109
6.3.3 Preliminary Work and Apparatus.....	110
6.3.3.1 The Training Program.....	110
6.3.3.2 The Chat Room Simulation.....	110
6.3.3.3 The Distracter Task.....	111
6.3.3.4 The Testing Software.....	112
6.3.3.5 The Laboratory and Hardware.....	112
6.3.3.6 Generation of the Chat Transcript.....	112
6.4 Results.....	113
6.4.1 Accuracy.....	113
6.4.2 Overall Comparison of RTs.....	114
6.4.3 Individual Comparison of RTs.....	115
6.5 Summary and Discussion.....	118

## **Chapter 7: Investigative Interviewing with QS-CMC**

7.1 Background.....	121
7.1.1 Interviewing through CMC.....	122
7.1.1.1 A Checklist.....	122
7.1.1.2 Interaction Management and Rapport.....	123
7.1.1.3 Questioning.....	124
7.1.1.4 Interview Length.....	124
7.1.1.5 Anxiety.....	125
7.1.1.6 Technological Facilitation.....	125
7.1.1.7 Summary.....	125
7.2 Aims of the Study.....	126
7.3 Deriving an Interviewing Strategy.....	126
7.3.1 Phase 1: Greeting and Rapport Building.....	126

7.3.2 Phase 2: Explaining the Interview Aims.....	128
7.3.3 Phase 3: Initiating a Free Narrative.....	128
7.3.4 Phase 4: Questioning.....	129
7.3.5 Phase 5: Varied and Extensive Retrieval.....	130
7.3.6 Phase 6: Summary.....	130
7.3.7 Phase 7: Closure.....	131
7.3.8 A Summary of the Explorative Themes.....	131
7.4 Method.....	132
7.5 Analysis.....	132
7.5.1 Free Narrative.....	132
7.5.2 Questioning.....	135
7.5.3 Varied Retrieval.....	137
7.5.4 Summation and Closure.....	138
7.5.5 Misalignment and Repair.....	138
7.6 Summary and Discussion.....	140
7.6.1 A Summary of the Findings by Theme.....	140
7.6.1.1 How Effective is the Free Narrative Phase?.....	140
7.6.1.2 Paralanguage, Empathy, and Listening.....	140
7.6.1.3 Avoiding Interruptions.....	140
7.6.1.4 Expressing the Interview Aims and Phase Instructions.....	141
7.6.1.5 The Transition into Mental Reinstatement.....	141
7.6.1.6 How Much Time is Spent Reinstating the Context?.....	141
7.6.1.7 How Effective is the Questioning Phase?.....	141
7.6.1.8 Resolution of Misunderstandings.....	141
7.6.1.9 How is the Message History Used?.....	142
7.6.1.10 How Successful were the Varied Retrieval Methods?.....	142
7.6.1.11 How Effective is the Summation Stage?.....	142
7.6.2 Is the Adapted ECI Suitable for SVA?.....	142
7.6.3 Improvements to the CMC-based ECI.....	143

## **Chapter 8: A Statement Validity Assessment within CMC**

8.1 Background.....	145
8.1.1 The Interview.....	147
8.1.2 The Content Criteria.....	148
8.1.3 The Nonstrategic Leakage Criteria.....	149
8.2 Hypotheses.....	149
8.3 Method.....	150
8.3.1 Participants.....	150
8.3.2 Design and Protocol.....	150

8.3.2.1 Brief.....	150
8.3.2.2 The Interviewing Software.....	151
8.3.2.3 The Interview Schedule.....	152
8.3.2.4 Debrief.....	154
8.3.3 Transcript Coding.....	155
8.4 Results.....	155
8.4.1 Tactics of Deceit.....	155
8.4.2 CBCA Scores.....	156
8.4.2.1 Overall Judgements of Veracity.....	156
8.4.2.2 Inter-rater Reliability.....	157
8.4.2.3 Multivariate Analysis of Individual Criteria.....	159
8.4.3 Deletions.....	160
8.5 Summary and Discussion.....	160
 <b><u>Chapter 9: Evaluation</u></b>	
9.1 Interpersonal Deceit within CMC: A Survey.....	164
9.2 RT Measures to Assess Guilty Knowledge.....	164
9.3 A CMC-Based SVA.....	167
9.4 Revisiting the Research Questions.....	168
9.5 Future Research.....	171
9.6 Conclusion.....	172
 <b>Appendix I: Addendum for Chapter 5.....</b>	<b>175</b>
 <b>Appendix II: Addendum for Chapter 6.....</b>	<b>180</b>
 <b>Appendix III: Addendum for Chapter 7.....</b>	<b>190</b>
 <b>Appendix IV: Addendum for Chapter 8.....</b>	<b>192</b>
 <b>Glossary.....</b>	<b>196</b>
 <b>References.....</b>	<b>201</b>

---

# List of Figures

Figure 1.1: A Schematic of IDT.....	13
Figure 4.1: The SIDE Model.....	62
Figure 6.1: Experimental Conditions used in (Farwell and Donchin, 1991).....	105
Figure 6.2: The Block Puzzle.....	113
Figure 6.3: Accuracy (%) of Responses.....	114
Figure 6.4: Correct-Response Times (ms).....	115
Figure 6.5: Box-Plots and Distributions of RTs.....	116
Figure 8.1: The Save Data and Ground Truth Dialogues.....	151
Figure 8.2: The Interviewer Assistant.....	152
Figure 8.3: Plot of Degree Against Confidence.....	157
Figure I-i: Questions and Instructions Used in the Analysis (Section 1).....	175
Figure I-ii: Questions and Instructions Used in the Analysis (Sections 2 and 3).....	176
Figure I-iii: The Introduction Page.....	177
Figure I-iv: The Instruction Page.....	177
Figure I-v: The Questionnaire Page.....	178
Figure I-vi: The About Page.....	178
Figure I-vii: The Exit Page.....	179
Figure II-i: Participant Information Sheet (Part 1).....	181
Figure II-ii: Participant Information Sheet (Part 2).....	182
Figure II-iii: Participant Information Sheet (Part 3).....	183
Figure II-iv: Participant Information Sheet (Part 4).....	184
Figure II-v: The Distracter Task.....	185
Figure II-vi: The Tester Software.....	186
Figure II-vii: The Training Software.....	186
Figure II-viii: The Chat Room Simulator.....	187
Figure II-ix: Sample of the Transcript used in the Simulation.....	188
Figure II-x: Instructions for the Block Puzzle Task.....	189
Figure III-i: The Chat Client.....	190
Figure III-ii: The Login Dialogue.....	190
Figure III-iii: Participant Information Sheet.....	191
Figure IV-i: Participant Information Sheet (Part 1).....	194
Figure IV-ii: Participant Information Sheet (Part 2).....	195



---

# List of Tables

Table 5.1: Context Interactivity and IDT.....	83
Table 6.1: Accuracy (%) of Responses Across Participants.....	114
Table 6.2: Correct-Response Times (ms) Across Participants.....	115
Table 6.3: Mean Response Times.....	117
Table 6.4: A Comparison between Guilty Probes and Pooled Data.....	117
Table 6.5: A Comparison between Innocent Probes and Pooled Data.....	118
Table 8.1: Degree of Veracity and Confidence Ratings.....	157
Table 8.2: Inter-rater Reliability.....	158
Table 8.3: Univariate Effects of the Subjective Criteria-Presence Scores.....	159
Table 8.4: Univariate Analysis of the Frequency Scores for Rater 1.....	159
Table 8.5: Univariate Analysis of the Frequency Scores for Rater 2.....	160
Table 9.1: Ordering and Gender Effects.....	165
Table II-i: Stimulus Word-Pairs used for the Guilty-Knowledge Test.....	180
Table IV-i: Deletion Statistics (Part 1).....	192
Table IV-ii: Deletion Statistics (Part 2).....	192

# Chapter 1

## Introduction

---

### 1.1 The Problem

Over the past ten years, the enthusiasm behind the proliferation of computer-mediated chat rooms has been tempered with reports of online deceit. The Internet, once a haven for the vulnerable and marginalised, has become a realm where the prudent must guard themselves from the malicious attacks of opportunists. Law-enforcement personnel are faced with the difficult task of policing a vast domain in order to exhume the deceptive practices that underlie illegal activity. Similarly, the proprietors of online communities are forced to expend resources moderating and regulating their forums as the only way to preserve their reputability. At present, both these tasks are performed in the absence of any substantive empirical evidence concerning the behaviours of online deceivers, and the clues that could be used to betray them.

An investigation of the way people deceive using computer-mediated communication (CMC) has another application. The role of lie-detection in law-enforcement primarily involves assessing the veridicality of a testimony from an eyewitness or suspect. According to Undeutsch (1989), when judging crimes of child sexual abuse, the child's statement is often the pivotal piece of evidence that determines the outcome of the allegation. Therefore, the ability to assess the veridicality of this statement is also pivotal. Such judgments cannot and should not be based on 'intuitive' assessments and biases founded upon a misinformed interpretation of nonverbal behaviour. Contemporary techniques that employ lie-detection, examine the verbal characteristics of a statement alongside other nonverbal and physiological clues. The success of these techniques often hinge upon the appropriate unbiased extraction of a testimony, and the appropriate use of diagnostic clues to behaviour (Vrij, 2000). It is suggested that the use of CMC, under some circumstances can provide a means of enhancing this control. Additionally it is believed that the use of CMC may provide a range of novel behavioural cues that reveal differences in the construction of both truthful and fabricated statements.

Throughout this thesis, these two problems have been addressed through the adaptation and synthesis of CMC and lie-detection research. This introduction will begin with a presentation of these two distinct research areas before the precise aims of the research detailed within this thesis are stated. At the end of this introductory chapter, the structure of the thesis is outlined along with some notes concerning its presentation.



## 1.2 Computer-Mediated Communication

Text-based CMC is a relatively new means by which humans can interact. The term can be used to refer to a variety of technologies such as email, mobile phone text-messaging, and computer-mediated chat. This form of communication is novel and affords many differences to traditional styles of collocation. The most salient of these, is the freedom from the body and an occlusion of the traditional nonverbal and paralinguistic cues to social interaction. In itself, this affordance has prompted the development of many online groups that are founded on the principles of anonymous interaction, and that are supposedly free from the shackles of stereotypes and marginalisation. However, this new form of anonymous communication comes hand-in-hand with the dangers of deceit. Such dangers stem from the relatively recent realisation that anonymous voices from an ethereal domain can have grave real world consequences, ranging from emotional hurt to the more sinister practices of paedophilic grooming.

The mid-1970s saw the emergence of social research investigating the organisational utility of embryonic CMC-based groups. The early research in this field was initiated by technically minded researchers who were interested in applying CMC within the workplace. Over the years, these investigations have passed over to social scientists from a number of different disciplines and subsequently the nature of CMC research branched away from technological determinism towards a more constructivist and relational style of thinking. Contemporary CMC research characterises the medium as a domain of opportunities that collocutors can leverage in pursuit of their varied and complex motives. Interaction within the medium is affected by many factors unique to this mode of discourse. Conversations are structurally re-engineered, anonymity can give rise to de-individuation effects, and identities are re-explored. Interpersonal relationships once thought to be difficult to maintain within the domain have been found to flourish, and even to exceed the intimacy of their offline equivalents.

## 1.3 Lie-Detection

The illusory panacea of accurate lie-detection has long been established in the consciousness of society. Scientific research in the field has emerged more recently, dispelling myths surrounding the diagnostic ability of some nonverbal behaviours and ascribing to the polygraph the low amount of internal validity that had been wrought out of wishful thinking and complacency. Contemporary techniques have begun to emerge, driven by sound physiological and psychological evidence. These include techniques that can assess veracity through the identification of verbal characteristics. Modern lie-detection and truth-validation techniques have attained modest but nevertheless significant accuracy levels through the careful application of systematic research. Coupled with the existing research on the communication of deceit, this can now provide a more accurate account of the psychological underpinnings of deceptive behaviour. It is this body of knowledge that allows the application of lie-detection research to the CMC domain.

## 1.4 Research Questions

The research detailed within this thesis investigates interpersonal deceit and lie-detection using computer-mediated communication (CMC). This field of inquiry is located within two distinct disciplines of study. These disciplines concern the most basic human practices of communication and deceit within the more contemporary contexts of technologically facilitated conversation and lie-detection. Within this thesis, these two fields will be synergised in a manner that is mutually beneficial to each, in order to answer two main questions. The first pertains to the exploration of naturalistic lie-detection.

### **How do deceivers perform their interpersonal deception strategies whilst communicating through computer-based quasi-synchronous conversation?**

This question has been investigated through the following objectives:

1. To build upon a theoretical synthesis between Interpersonal Deception Theory and CMC theory that was proposed by Burgoon et al. (2003), and to extend this application to the field of quasi-synchronous chat.
2. To attain a broad range of qualitative accounts of online deceptive episodes in order to identify tactics used by deceivers, attitudes towards CMC deceit, and to research the key behaviours that underlie suspicion.

The second broad question examined within this thesis concerns the novel use of CMC as a lie-detection tool.

### **Is it possible to create a lie-detection technique that can distinguish deceptive behaviour reliably from the behaviour of truth-tellers in a controlled CMC-based setting?**

This question has been investigated through the following objectives:

3. To investigate the ability to detect low-level response data, using standard computer equipment, in order to reveal guilty knowledge of a CMC-based disturbance.
4. To investigate how CMC can be used to add more control to the process of attaining an eyewitness testimony for subsequent codification through a criteria-based truth-validation technique.
5. To investigate the effectiveness of Criteria-based Content Analysis in distinguishing truthful and fabricated stories generated through CMC.
6. To investigate the effectiveness of certain nonverbal lie-detection criteria (derived as part of Objective 1) in distinguishing truthful and fabricated stories generated over CMC.

## 1.5 Thesis Overview

The literature review for this thesis is contained within the next three chapters. Chapter 2 explores the ways in which writers have conceptualised and categorised different types of lying and deceptive behaviour. This is coupled with a review of Interpersonal Deception Theory that describes the performance of deceptive episodes between collocutors. In Chapter 3, the scientific literature on lie-detection is reviewed, concentrating on those areas that are the most applicable to the domain of CMC. Chapter 4 then details the dominant theories of mediated communication that have been used within the field of CMC research along with a discussion of internet deception and the construction of the 'self' online.

Chapter 5 begins with a theoretical account exploring the first two objectives set within this thesis. The theories of CMC, interpersonal deceit, and lie-detection are brought together in a manner that extends and expands upon the only other existing attempt to achieve this goal. To this end, the propositions of Interpersonal Deception Theory (Buller and Burgoon, 1996) and its recent adaptation to CMC (Burgoon et al., 2003) are re-explored, and expanded-upon in order to more fully encapsulate naturalistic deceptive behaviour. As a part of this exploration, a set of quasi-synchronous CMC-based nonstrategic leakage clues to deceit are proposed for the first time, derived from factors that influence similar behaviour in face-to-face encounters.

The need for research exploring the strategic behaviours of deceivers is addressed in the latter part of Chapter 5. This empirical work was accomplished using a qualitative survey under a constructivist paradigm and examined how people reported their experiences of deceptive episodes. The survey was constructed as an interactive website and was made accessible to the public. The analysis of the questionnaire has uncovered some of the tactics and behaviours used by both deceivers and the deceived.

The next three chapters pertain to the other objectives set within this thesis. Chapter 6 describes research that applies a lie-detection technique based upon the 'guilty knowledge' approach. This experimental study is the first reported that successfully uses standard mouse-click response times as a dependant variable for detecting deceit. The scenario used within this study also explored whether such a technique could be used to detect a novel form of chat room based mock crime. This study provides further evidence to suggest that various other 'nonverbal' CMC-based clues could await exploration.

Chapters 7 and 8 concern the use of CMC for investigative interviewing within a popular truth-validation technique called Statement Validity Assessment (SVA). Two studies examined the effectiveness of using CMC as a means for interviewing eyewitnesses and judging the veridicality of their statements. Twenty-five truthful and twenty-five fabricated statements (generated using CMC) were tested using a content analysis technique for assessing veracity. In addition, low-level nonstrategic information about the behaviour of the participants was recorded.

The evaluation of the research described within this thesis has been presented in Chapter 9. This evaluation focuses on issues of validity, realism, and reliability that were omitted from each individual research chapter. This final chapter is concluded with a discussion of how the research conducted within this thesis relates to the objectives posed within this introduction, before the opportunities for further exploration are discussed.

More information relating to the four empirical studies described in Chapters 5, 6, 7, and 8 are found in the four respective appendices at the back of this thesis.

## **1.6 Notes**

Within this thesis, unless otherwise indicated the term 'lie-detection' is taken to encompass both lie-detection and truth-validation techniques. The terms false-positive and false-negative have been used throughout, and the reader is alerted to the fact that the interpretation of these two terms will differ according to whether the discussion is referring to a lie-detection or a truth-validation technique. The use of the acronym CMC, unless otherwise indicated refers to text-based CMC.

## Chapter 2

# Interpersonal Deception Theory and the Semantics of Deceit

*“... and let's have no displays of indignation. You may not have known but you certainly had suspicions. If we've told lies, you've told half-lies, and a man who tells lies, like me, merely hides the truth, but a man who tells half-lies has forgotten where he put it.”\**

---

The reliable and valid detection of deceit is a precarious practice. As the communicative skills of human beings surpassed those of the animal kingdom, so too did our abilities to fashion deceptive signals. Whereas an animal might use a deceptive signal to ward off predators or to trick prey, so too have humans adopted deceptive practices in order to facilitate their more varied, complex, and interpersonal motives. Deception, in all its forms has become prevalent in our society, and as new forms of communication develop, so too does the means by which we can become deceptively engaged.

One of the main challenges for lie-detection, as with any other measuring or detection device, is the appropriate classification of the phenomena under scrutiny. The subjectivity surrounding the interpretation of 'untruths' in everyday life can make this a particularly difficult task. Nonetheless, it is absolutely essential for any detection tool to have a grasp of what it is that is being measured, and to be able to accurately distinguish that phenomenon from its neighbours. The purpose of this chapter is to provide a review of how scientists and philosophers have attempted to define and classify the phenomena of 'lying' and 'deceit'. This will provide a clearer framework for which to understand the use of these terms throughout this thesis. To support this grounding, the work on Interpersonal Deception Theory (IDT) is presented, a bold attempt to organise research describing the nature of deceptive interpersonal episodes. In sum, this chapter will not attempt to address the psychology of why people decide to lie, but will focus instead on how they go about it, and the type of information they attempt to lie about.

---

\* Quote taken from the motion picture "Lawrence of Arabia" (Columbia Pictures, 1962). Here Lawrence is being informed of the British Government's intention to retain control of Arabia following the war, despite previous assurances to the contrary.

## 2.1 Searching for a Definition of a Lie

Commonly used terms such as 'lying' and 'deceit' are wide-ranging semantic constructs that can encompass a variety of phenomena connected to a 'reversal of truth'. When attempting to provide their definitions, it is important not to nullify the utility of these terms by making them too broad. Barnes (1994) makes this criticism of Bailey's (1991) use of the word 'lying', suggesting that an over enthusiastic use of the term can render it an analytically blunt tool. The problematic everyday semantics of lying can itself serve as a vehicle to deceive. It has been said that, "[s]uccess in telling lies depends in large part on exploiting the possibilities inherent in natural languages" (Barnes, 1994; p.115). Such possibilities are classically exemplified with the many formulations of the 'liar's paradox'. The statement 'This sentence is false' exemplifies where natural language can be ambiguous when describing logical notions of truth and falsity. Therefore, it is important to find a definition that is both narrow enough to be useful, and that is acute enough to encapsulate the essence of lying.

The ways in which the definitions of 'lying' and 'deceit' are construed depends essentially on focus. For example, definitions can be driven by moral or religious concerns. A pious individual may redefine the term 'lie' in order to adhere to an impractical orthodoxy that would otherwise proscribe a variety of activities such as 'white lies' or ironic humour (Bok, 1999). Sissela Bok (1999) uses a wider definition in order to neutralise the troublesome morality concerns that have a tendency to be smuggled into narrower definitions. Bok defines a lie as "an intentionally deceptive message in the form of a statement" (p.15) and this provides a good starting point. Here the key characteristic of intentionality is stressed, and 'lying' has been disambiguated from the broader category of 'deceit' by classifying it as a deceptive *statement*.

Bok acknowledges that this definition encompasses a broad spectrum of subtle variants of lying and subsequently focuses the rest of her book on 'clear-cut lies'. These are lies "where the intention to mislead is obvious, where the liar knows what he is communicating is not what he believes, and where he has not deluded himself into believing his own deceptions" (Bok, 1999; p.16). The last clause excludes cases of self-deception.

Bailey (1991) finds Bok's definition of the 'clear-cut lie' to lack utility from the standpoint of practical ethics. The definition seems to rely upon an understanding of "the inner workings of other peoples' minds" (p.9) and so the clear-cut lie becomes difficult to identify. Instead, Bailey utilises a broader category of 'untruths' as a term without moral standing. Untruths come in three forms, deceit, error and fiction. Lies would fall under the broad category of deceit (the analytically blunt tool discouraged by Barnes).

Barnes (1994) adopts Bok's definition, but expands it to include self-deception<sup>1</sup>. In addition, he stresses that the definition should ignore the success or the failure of the attempted deceit and that the classification should be based on intent. Furthermore, he agrees with Bok's distinction

---

<sup>1</sup> Barnes defines self-deceit as the "process whereby there is an inner dialogue in which one segment of a personality deceives another segment, and where the lying segment remains aware of the deceit" (Barnes, 1994; p.89).



between questions of morality and those of epistemology. It is the liar's *beliefs* surrounding the truth that are important.

Ekman (1985) chooses to use the terms lying and deceit interchangeably. He adopts the broad definition of "a deliberate choice to mislead a target without giving any notification of the intent to do so" (p.41). Here another important component of the semantic map is revealed, that of prior notification. Bok (1999) reveals that such prior notification can be signaled by both verbal and nonverbal cues, but the target must not be forewarned that they are going to be lied to (as would usually be the case in a stage performance such as a magic show). Ekman's definition provides the interpretation of 'deceit' that is used within this thesis.

Vrij's (2000) definition of lying incorporates many of the components highlighted by these authors. He defines deceit as "a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue" (p.6). This is a better working definition, but there are still some interesting scenarios of deceit that it can incorporate, especially concerning the mechanisms the liar uses to attempt to instill the untrue belief. For example, a well-known liar may tell the truth expecting not to be believed and thus by attempting to instill an untrue belief, they could still be considered to be lying. Vrij's definition of deceit shall be used throughout this thesis as the definition of 'lying', with the modification that the 'attempt' must be a 'verbal attempt'.

## **2.2 Categorising Lies**

The definitions used by Ekman and Vrij cover a wide variety of phenomena that are of interest to a lie-detection researcher. A full categorisation of deceit would be a challenging task, incorporating those techniques adopted by both animals and humans across a wide variety of signalling media. Here, the focus shall be on the attempts to categorise lying which itself has had a long history within the literature.

### **2.2.1 Criteria of Categorisation**

Attempts to categorise lying phenomena have been founded upon the theoretical focus and the particular definitions used by their originators. These typologies usually involve one or more of the following dimensions.

#### **2.2.1.1 Lie Complexity**

Lies may be categorised according to their level of complexity. Complexity may be assessed for the lie per se, or for the larger deceit in which the lie is a component. A complex lie could involve sophisticated techniques such as the 'incorrect-inference dodge' (Ekman, 2001)<sup>2</sup>. In this scenario, the deceiver presents facts in such a way that the dupe is led into making false assumptions or inferences. This form of lying is complex because in order to be successful the

---

<sup>2</sup> This has also been referred to as a 'virtual' lie (Black, 1983)

deceiver is required to correctly judge exactly how to trick the dupe into making the false inference. Another example of a complex lie is what Barnes (1994) refers to as 'lying by proxy'. Here, a third party is tricked into relaying the deceitful information to the dupe whilst at the same time believing in the veridicality of the message. Connivance can also be seen as a complex form of lying, whereby a person gives no indication that they are aware that the statement addressed towards them is a lie.

### **2.2.1.2 Outcome**

Lies may be categorised by the consequences that occur because of their action. For example, St Augustine classified lies into the following eight categories (Bok, 1999).

1. *lies that involve religion*
2. *lies that injure someone unjustly*
3. *lies that benefit someone whilst harming another (without 'physical defilement')*
4. *lies told for the pleasure of deceiving*
5. *lies told from a desire to please others in smooth discourse*
6. *lies that harm no one and benefits some person*
7. *lies that harm no one and benefits someone excluding where a judge is questioning*
8. *lies that harm no one and benefits someone to the extent that it protects someone from physical defilement*

This categorisation deals primarily with who is harmed, the type of harm, and who benefits from the episode. Lies can also be evaluated as either high-stake or low-stake depending upon the severity of the repercussions if the lie is uncovered. This distinction (along with complexity classifications) is often used within lie-detection studies that are founded upon nonverbal clues (see Ekman, 2001). This is because it is sometimes easier to detect deceit under these criteria when the deceiver is placed under pressure (Ekman, 2001).

### **2.2.1.3 Moral Turpitude**

Barnes (1994) classifies lies as being either benevolent or malicious, and similarly Goffman (1975) refers to fabrications that are either benign or exploitative. Here, the motivation of the deceiver serves as the distinguishing criteria. Bailey refers to these motivations and the moral evaluations associated with them as the *turpitude dimension* of lying. Thomas Aquinas (cited in Bok, 1999) similarly classifies lies as being either 'officious', 'jocose', or 'mischievous'. However, he suggests that this classification may not be preferable from a philosophical standpoint as "a division should be made according to that which pertains to a thing by reason of its nature [...] the intention of the effect resulting from a moral act is something beside and accidental to the species of that act, so that an indefinite number of effects can result from one act" (Bok, 1999; p.255). This point is particularly relevant for the practice of lie-detection, where the intention of the deceiver within a moral context might very well be incidental to the mechanisms used to construct a lie and consequently incidental to the process of detecting it.

#### **2.2.1.4 Cultural Factors**

Classifications based upon moral turpitude are naturally connected to the culture within which these systems reside. Barnes (1994) claims that, there is no cross-cultural uniformity of when lies are malevolent and when they are benign. For the Navajo, a lie only becomes deceitful the fourth time it is uttered. In Lebanon, villagers have been found to have quite a playful attitude to lying and in rural areas of Greece, cheating and deceiving within financial deals can be considered quite virtuous (Barnes, 1994).

#### **2.2.1.5 Lying Tactics**

Many authors have separated lies into categories that reflect the mechanisms used to induce the deceitful belief. A common distinction is between lies that relate false information, and those that omit information. Douglas (1976) refers to these as lies of commission and lies of omission respectively. Thomas Aquinas (cited in Bok, 1999) argued that this is the primary way that lies should be categorised. He segregates lies into 'boasting' (lies that go beyond truth) and 'irony' (lies that fall short of truth). Ekman (1991) also examines these two modes of deceit (referring to them as 'falsification' and 'concealment'), but also mentions 'misdirecting', 'telling the truth falsely', 'half concealment', and the 'incorrect inference dodge' mentioned earlier. DePaulo and her colleagues (1996a) have used the categories 'outright lies', 'exaggerations', and 'subtle lies' (subtle lies refer to 'literal truths designed to mislead'). These extra categories are beneficial where a broader definition of lying is used. The simpler categorisation loses its utility when dealing with statements that are both commissive and omissive.

#### **2.2.1.6 Actors**

Lies may be classified according to the type of parties involved. For example, there are paternalistic lies by adults to children, lying to liars or enemies, or therapeutic lies to the sick (see chapters in Bok, 1999). Lies may also be separated according to the number of actors involved, distinguishing collusive lies (many-to-one), dyadic lies (one-to-one), basic lies (one-to-many), and self-deceit. There are also some who would wish to distinguish pathological lying (*pseudologica fantastica*) depending on the author's views towards the validity of the condition.

Lies to people of a different authority are commonplace (Hartung, 1988 uses the terms 'deceiving down' and 'deceiving up'). Bailey (1991) observes that in many cases, lies are 'inextricably entangled with power' and he labels lies from authority as *hegemonic* lies<sup>3</sup>. Plato's 'noble lie' is an example of hegemonic lie. Plato (Waterfield, 1998) argued that it was justifiable for leaders to lie to the populous if ultimately it was for their own benefit.

---

<sup>3</sup> Bailey makes some interesting comments concerning hegemonic lies. He asserts that many hegemonic lies are actually protected by those champions for truth in our society. For example, investigative journalists attack individuals but leave the system intact. This creates the illusion of an open society and of a governing class responsive to those below it (p.91), but nevertheless, these journalists still serve to maintain the hegemonic, basic lie by "preventing the build up of excessive friction" (p.82).

### 2.2.1.7 Motivation

Lies can be characterised according to the specific motivations of the liar. Vrij (2000) provides some motivational categories.

1. *lying to make a positive impression*
2. *lying in order to take advantage*
3. *lying to avoid punishment*
4. *lying for another's benefit*
5. *lying for the sake of social relationships*

In some situations, a deceiver's motivations may be clear. In a forensic polygraph test for example, the primary motivation would seem to be to avoid punishment. However, other possible reasons for the lie might be to benefit another person or even to take advantage. Such complexities have meant that lie-detection investigators have tended to avoid detection strategies that are based upon these distinctions.

## 2.2.2 Categorising the Lies of Children

The study of children's lies is a pedagogical concern. The first issue of the journal *Pedagogical Seminary*<sup>4</sup> contained a categorisation of children's lies that were obtained through group discussions with three hundred children at Kindergartens across Boston (Hall, 1891). This categorisation is repeated here, and reflects many of the facets presented in the previous section.

1. *Pseudophobia*
2. *The Lie Heroic*
3. *Lies for Enemies*
4. *Selfishness*
5. *Imagination and Play*
6. *Pseudomania*
7. *Pantomimed Lies*

Pseudophobic lies are characterised by an incredible sense of fear associated with the deviation from veracity. In contrast, Pseudomania represents the very joy of lying, the exhilaration of deceit and is synonymous with *pseudologica fantastica*. The Lie Heroic represents a form of self-sacrificial lying, such as lying in order to save a friend's life. Pantomimed lies use expressions that palliate deceit and can involve inaudible and unspoken statements that further qualify or even contradict statements that are spoken aloud (Bok, 1999 refers to these as 'mental reservations').

Modern classifications of the lies of children are based upon the developmental cognitive abilities that are necessary for lying in the first place; abilities that were poorly understood one

---

<sup>4</sup> Later renamed the Journal of Genetic Psychology

hundred years ago. Vasek (1986) defines three necessary capacities for lying including the ability to perceive intent, cognitive 'perspective-taking', and communication skills. Piaget (1932) found that the notions of 'lying' behaviour at a very early age were synonymous with the telling of a 'naughty word'. Between the ages of six and ten, the child would refine this notion with respect to the truth-value of the lie (considered as the telling of any statement that is *false*). Piaget found that many children aged between six or seven would evaluate the severity of the lie according to a quantitative notion of falsity. Therefore, the severity of the lie was often measured "by the degree of its incredibility to adults" (p.147) and Piaget describes this perspective as being a form of moral realism. Only until the age of ten or eleven does the child begin to concern itself with notions of intentionality. The child is 'liberated' from moral realism as they progress towards a mutual form of respect.

## 2.3 Interpersonal Deception Theory

Interpersonal Deception Theory (IDT) is the attempt by David Buller and Judee Burgoon to produce a communication theory of deception. Their approach emphasised the interpersonal, mutualist nature of deception that had previously been neglected by researchers (Buller and Burgoon, 1996a). The theory was presented and critically appraised in a special issue of 'Communication Theory' (August 1996) and the description that follows has largely been based upon the issues presented there.

### 2.3.1 The Theory

In 1981, Zuckerman, DePaulo, and Rosenthal published a seminal article that reviewed previous research on the verbal and nonverbal communication of deception (Zuckerman et al., 1981). The authors understood that because deception was not an affective response, it was unlikely to be associated with any particular kind of nonverbal expression. Despite this observation, they presented four factors underlying behaviour that could *correlate* with deception. These included (1) the deceiver's attempts to control their behaviour, (2) autonomic physiological arousal, (3) affective responses from duping delight or detection apprehension, (4) and increases in cognitive load resulting from the complexity of constructing a lie.

This perspective can be referred to as a 'leakage approach', whereby the clues betraying deceptive behaviour have somehow 'escaped' the control of the deceiver. The empirical research investigating these clues is often conducted within the laboratory. Typically, the participant is asked to tell a lie in front of a video camera, under different experimental conditions, whilst their responses are filmed for subsequent codification. Whilst this research is largely consistent with the phenomena predicted by the four-factor approach, it underplayed the importance of the mutual interpersonal processes that exist within everyday interaction. This shortcoming was addressed by Buller and Burgoon (1996a) in their tentative formulation of IDT.

The problem with previous research was that it eschewed (intentionally or otherwise) the processes of interpersonal interaction and their effects upon behaviour. In order to understand the behavioural displays of deceptive interpersonal activity, it is important to not only understand the non-strategic leakages predicted by the four factor approach, but also to understand the strategic ways that deceivers behave in order to appear convincing. Previous research had been conducted using relatively un-interactive scenarios and Buller and Burgoon predicted that as interaction increased this would affect changes in (1) the goals of the deceiver, (2) communication skills, (3) mutual influence, (4) the potential for feedback, (5) and the inherent dynamics of interpersonal communication. They surmised that within interactive contexts, the non-strategic leakage activity predicted by the four-factor approach would be mediated by the superordinate displays of the mutual and strategic activity predicted by IDT.

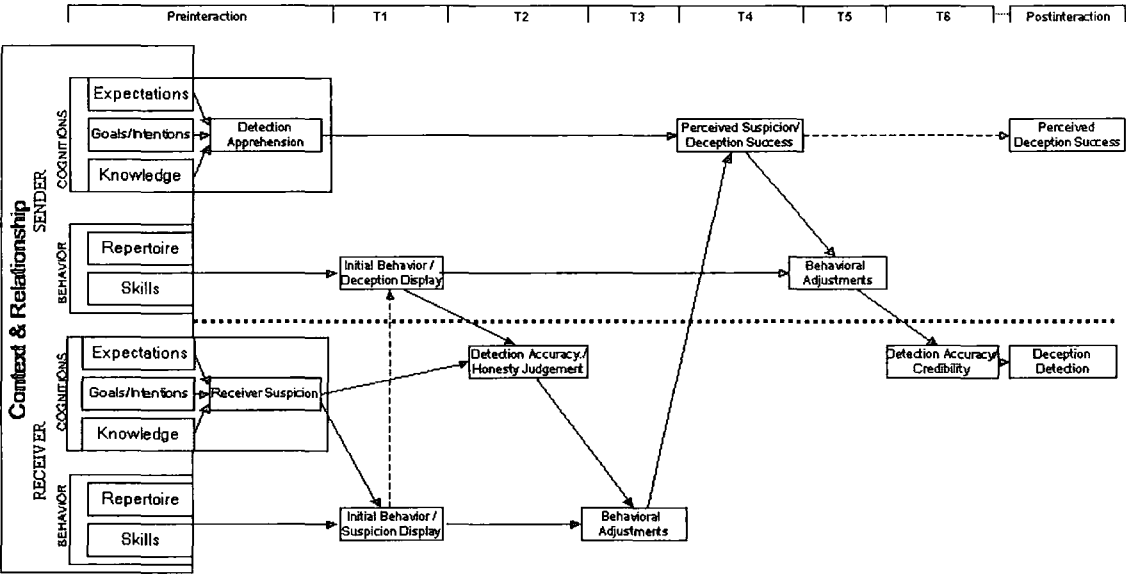


Figure 1.1: A Schematic of IDT (adapted from Buller and Burgoon, 1996a; p.211)

Buller and Burgoon provided a schematic in order to clarify the processes of strategic activity by both a sender and a receiver during a deceptive exchange; this has been reproduced in Figure 1.1 above. The sender and receiver both come to the interaction with certain behavioral norms and skills that typify their interaction. Additionally, they bring certain expectations, intentions, and knowledge to the interaction that act as a form of pre-interaction bias. For the sender these preconceptions will affect their initial apprehension at being detected, for the receiver they will affect initial suspicion. During the interaction, co-interactants display behaviors that will mutually lead to behavioural adjustment, including sender adjustment to receiver displays of suspicion. The process is iterative and functions may occur simultaneously (the time scale on the schematic is a simplification). Following the interaction, the sender will be left with a perception of deception success, and the receiver with a perception regarding the likelihood of deceit.

From certain assumptions regarding interpersonal and deceptive behaviour, IDT is founded upon a set of propositions from which testable hypothesis might be derived. These propositions are discussed briefly below <sup>5</sup>.

### **2.3.1.1 The Propositions of IDT**

Deceptive displays vary according to the degree of interactivity and the nature of the existing interpersonal relationship between the sender and receiver. This is referred to as the *interactivity principle* of IDT.

1. *Sender and Receiver cognitions and behaviors vary systematically as deceptive communication contexts vary in (a) access to social cues, (b) immediacy, (c) relational engagement, (d) conversational demands, and (e) spontaneity.*
2. *During deceptive interchanges, sender and receiver cognitions and behaviors vary systematically as relationships vary in (a) relational familiarity (including informational and behavioral familiarity) and (b) relational valence.*

Behaviour during a deceptive display is a product of both strategic and non-strategic action, and this is more prevalent than it would be during everyday veridical behaviour.

3. *Compared with truth tellers, deceivers (a) engage in greater strategic activity designed to manage information, behavior, and image and (b) display more non-strategic arousal cues, negative and dampened effect, non-involvement, and performance decrements.*

Receivers will have presumptions that will affect their initial level of suspicion. Such suspicion can be detected by senders and this will subsequently affect their early levels of 'detection apprehension'. Over time, senders will increase their strategic displays within interactive contexts dampening the activity of non-strategic leakage.

4. *Context Interactivity moderates initial deception displays such that deception in increasingly interactive contexts results in (a) greater strategic activity (information, behavior, and image management) and (b) reduced non-strategic activity (arousal, negative or dampened affect, and performance decrements) over time relative to non interactive contexts.*
5. *Sender and receiver initial expectations for honesty are positively related to degree of context interactivity and positivity of relationship between sender and receiver.*
6. *Deceivers' initial detection apprehension and associated strategic activity are inversely related to expectations for honesty (which are themselves a function of context interactivity and relationship positivity).*

---

<sup>5</sup> The propositions in italics have been reproduced as per (Buller and Burgoon, 1996).

Those senders who deceive in order to benefit themselves will adopt a deceptive pattern that is highly strategic in keeping with those goals.

*7. Goals and motivations moderate strategic and non-strategic behavior displays*

*a) Senders deceiving for self-gain exhibit more strategic activity and non-strategic leakage than senders deceiving for other benefits.*

*b) Receivers' initial behavior patterns are a function of (a) their priorities among instrumental, relational, and identity objectives and (b) their initial intent to uncover deceit.*

As a receiver becomes accustomed to the sender's communicative style, this will be detected by the sender who will subsequently feel that they have to work harder in order to be seen as non-deceptive. This will result in an increase in both strategic and non-strategic behaviours. Experienced deceivers will be able to minimise their non-strategic behaviour and engage in more strategic behaviour. For example, the nonverbal leakage of stress, exhibited through an increase in adaptors (nonverbal gestures that manipulate another part of the body) could be minimised through the strategic behaviour of appearing more expressive (and therefore more confident).

*8. As receivers' informational, behavioral, and relational familiarity increase, deceivers not only (a) experience more detection apprehension and (b) exhibit more strategic information behavior, and image management but also (c) more non-strategic leakage behavior.*

*9. Skilled senders better convey a truthful demeanour by engaging in more strategic behavior and less non-strategic leakage than unskilled ones.*

Judgments concerning the veridicality of the sender are in many ways based on previously formed impressions and biases as well as the skills of the sender. Naturally, these influences can have the opposite affect on the accuracy of those judgments. A sender's deviation from their expected behavioural responses are thought to increase receiver suspicions, but also receiver detection accuracy as well. Suspicion is detectable by the deceiver, who will in turn try to adapt their behaviour to appear more credible

*10. Initial and ongoing receiver judgements of sender credibility are positively related to (a) receiver truth biases, (b) context interactivity, (c) and sender encoding skills; they are inversely related to (d) deviations of sender communication from expected patterns.*

*11. Initial and ongoing receiver detection accuracy are inversely related to (a) receiver truth biases, (b) context interactivity, (c) and sender encoding skill; they are positively related to (d) informational and behavioral familiarity, (e) receiver decoding skills, and (f) deviations of sender communication from expected patterns.*



12. *Receiver suspicion is manifested through a combination of strategic and non-strategic behavior.*
13. *Senders perceive suspicion when it is present.*
14. *Deviations from expected receiver behavior increase perceptions of suspicion.*
15. *Receiver behavior signalling disbelief, uncertainty, or the need for additional information increase sender perceptions of suspicion.*
16. *Suspicion (perceived or actual) increases senders' (a) strategic and (b) non-strategic behavior.*

Despite the presence of both strategic behaviours and non-strategic leakage, typical interpersonal behavioural dynamics are the superordinate processes in deceptive communication.

17. *Deception and suspicion displays change over time*
18. *Reciprocity is the predominant interaction adaptation pattern between senders and receivers during interpersonal deception.*

Perceived impressions of deception or detection success following an interaction are subject to a latency effect.

19. *Receiver detection accuracy, bias, and judgements of sender credibility following an interaction are a function of (a) terminal receiver cognitions (suspicion, truth biases), (b) receiver decoding skill, and (c) terminal sender behavioral displays.*
20. *Sender perceived deception success is a function of (a) terminal sender cognitions (perceived suspicion) and (b) terminal receiver behavioral displays.*

### **2.3.2 Critique**

IDT was presented in a special issue of 'Communication Theory' (August 1996). In the same issue, it was reviewed by some key researchers in the field of interpersonal deception including James Stiff, and Bella DePaulo and her colleagues. Their critique of IDT primarily surrounds the status of IDT as a practical theory of interpersonal deception and has been included here as being representative of the challenges facing IDT.

Stiff (1996) is keen to point out the merits of Buller and Burgoon's work. He applauds the contributions of IDT to the literature as a careful description of the behavioural displays that can lead to successful deceit or its detection. IDT also draws attention to the importance of temporality, mutuality, and pre-interaction in deceptive interchanges. However, the status of IDT as a theory (formative or otherwise) is in dispute. IDT has been criticised for failing to present a central causal mechanism that can be used to explain a wide variety of deceptive phenomena (Stiff, 1996; DePaulo et al., 1996b). Such a mechanism is necessary for any theory to have

predictive power, as DePaulo and her colleagues put it, there is no central, intriguing 'why' question in IDT. IDT is descriptive but not explanatory.

Burgoon and Buller (1996) reply to this critique by suggesting that the propositions merely look more descriptive than explanatory because of their high level of abstraction. They suggest that explanatory mechanisms do reside within the propositions, assumptions and their relationships.

"Among the primary explanatory mechanisms are the goals, skills, and communicative repertoire that lead actors to initiate certain behavioral patterns and eschew others; their enacted communication styles, which reflect deliberate (strategic) and inadvertent (non-strategic) activity; the inevitable sequencing and adaptation of behavior as goals are met or thwarted and interactants entrain to one another's communicative style; and the functional requirements and constraints of the communication situation itself" (Burgoon and Buller, 1996)

The authors warn against an excessive reductionism that would interpret their causal mechanisms as merely descriptive as opposed to explanatory. Burgoon and Buller defend their theory further, by stating that the 'requirement' of a single unifying causal mechanism reflects a "very narrow understanding of what a theory should look like" (p.315) and suggest that multiple mechanisms should be looked for and often are in existing scientific theories.

DePaulo and her colleagues (1996b) criticise IDT further by stating that many of the propositions are not falsifiable, and that they do not seem to be systematically related (a proposition may be added or removed with little consequence to the others). Burgoon and Buller (1996) address the falsifiability question by claiming that it is the specifically derived hypotheses and not the propositions themselves that are falsifiable. Further, they claim that the IDT propositions have not haphazardly been thrown into the theory, but instead represent the most relevant factors and relationships concerned with interpersonal deception.

Stiff (1996) continues his critique by claiming that Buller and Burgoon are poorly conceptualising the nature of interpersonal communication in IDT. He claims that the theory seems to pertain more to the *interactive* component of communication than the *interpersonal*. He suggests that there are areas of communication that are not covered by the theory, mainly the less interactive 'and' less interpersonal areas. This theoretical selectivity had arisen because of the poor definition of interpersonal communication that took face-to-face communication as the primary exemplar of the phenomenon. By narrowing their focus, Buller and Burgoon were limiting the explanatory scope of their theory. DePaulo and her colleagues (1996b) also question the superordinate role of strategic behaviours in deceptive displays. They do not believe that the strategic, self-reflexivity of co-interactants posited by IDT is actually as prevalent within interpersonal episodes as the theory assumes.

Burgoon and Buller (1996) do accept that their work has primarily involved highly interactive face-to-face communication, but they still maintain that IDT can be applied more widely

(including to the domain of computer-mediated communication). Specifically they adopt a model of interactivity whereby different interpersonal communication episodes lie along a continuum of interactivity.

Another question raised by Stiff (1996) was whether a new communication theory of deception was needed at all. Maybe the communication of deception can be sufficiently explained by existing theories (c.f. Stiff, 1995) and IDT does not have anything new to add. It could be argued that the central goal of deception is persuasion, and that theories of persuasion and social influence can sufficiently encompass deceptive episodes.

Broadly speaking, the main issues of IDT discussed in the special issue, surrounded the degree of abstraction in which a theory should lie in order to be predictively useful and whether there is a need for such a theory of interpersonal deception in the first place. Additionally, DePaulo and her colleagues attacked the empirical work used to exemplify IDT's utility (c.f. Burgoon et al., 1996a; Buller et al., 1996; in the same issue), as being overly complex and heavily confounded (DePaulo et al., 1996b). It is not disputed that Buller and Burgoon have provided a useful model with IDT; however back in 1996, it was unclear as to how much further the model could go. Subsequent empirical work has shown that IDT does indeed have more predictive and explorative power than was initially thought by its reviewers, and recently this approach has even been applied to the domain of computer-mediated communication. A discussion of the IDT research that examines CMC and the interactivity principle will be deferred until Chapter 5.

### **2.3.3 Empirical Support for IDT**

The main experimental study 'showcased' by Buller and Burgoon in the 1996 issue was criticised heavily by DePaulo and her colleagues (1996b). It was accused of being overly complicated and for containing confounds such as ordering effects and the underplaying of participant gender. In addition, the study included a vast number of experimental conditions (the authors counted ninety-six potential between-subject cells across only sixty-six participant dyads; DePaulo et al., 1996b). Burgoon and Buller (1996) responded to this critique by suggesting that the highly complex relationships investigated in their research was indicative of the complex nature of reality rather than a product of their design, and that not all of the results were published in the reports provided. Nonetheless, subsequent research has been far more focused with much clearer results.

In another article that same year, Burgoon and her colleagues (1996b) produced an essay that further elaborated upon the different perspectives of the sender, receiver and the observer. From a pilot study, they observed a high correlation (with large effect sizes) of gestalt perceptions regarding the veracity of the exchange between each of these different parties. The poor abilities of observers and receivers to detect deceit has been reported elsewhere (c.f. Ekman, 2001), however the authors had further clarified that the reasons for this lay in the joint construction of reality between the sender and the receiver rather than solely as a result of poor lie-detection on the part of the receiver.

Some of the recent work testing IDT has focused on the effects of conversational involvement (Burgoon et al., 1999; White and Burgoon, 2001). In a study that involved 122 participants interacting in dyadic question-answer sessions, Burgoon found that deceivers (especially socially-skilled ones) were generally more involved in their interaction. Encoding skills were found to be more critical in ensuring deceptive success than decoding skills, as was the involvement of the sender. Finally, they tested the effects of feedback and found that uninvolved receivers actually caused the deceivers to heighten and even overcompensate their strategic activities (this finding has interesting implications for lie-detection). White and Burgoon (2001) examined initial involvement in more detail with respect to Interaction Adaptation Theory (IAT; Burgoon et al., 1995). From IAT they categorise the three fundamental influences upon initial involvement as resulting from (in decreasing order of influence) requirements, expectations, and desires. Differences in the self-assessments of these factors will cause differences in the initial involvement patterns between truth-tellers and deceivers. For deceivers, the *requirements* to interact have associated negative effects and this can suppress initial involvement. The *expectation* to interact and the convention of social norms have a more expressive effect, as do *desires* pertaining to self-presentation. Contrastingly, the *desire* to control behaviour can act to suppress that behaviour.

## 2.4 Summary

This chapter has explored the semantics of deception and has chosen definitions for *lying* and *deceit* that are used throughout this thesis. These definitions are broad enough to appropriately encapsulate a variety of deceptive phenomena but also narrow enough to distinguish deceit from phenomena such as mistakes, illusions, self-deceit, or the simple conveyance of false information. Various schemes for classifying lies were presented, and it was observed that these schemes are dependant upon the focus of the activity to which they are applied. It would seem logical that the practice of lie-detection would require a classification scheme of its own, and the derivation of such a scheme has been deferred until Chapter 5.

In addition, the central propositions of Interpersonal Deception Theory were presented. IDT provides a framework of how people experience and interact in everyday deceptive episodes both strategically and non-strategically. It was shown that IDT may have some weaknesses in its theoretical status, but nonetheless has proven to be a useful framework for conceptualising everyday interpersonal deceptive encounters.

However, the practice of lie-detection usually does not rely upon scrutinising everyday interpersonal deceptive encounters. In fact, the direct manipulation of the interpersonal context (sometimes even to minimise interaction) is a key mechanism by which deceivers are forced to lie in a particular manner that makes lie-detection simpler. Naturalistic lie-detection has often been found to be too difficult partially because of the superordinate interpersonal processes that eschew nonverbal leakage, and partially because of the difficulty of distinguishing between the truthful and deceptive components of a communicative episode. This is not too say that

naturalistic lie-detection is impossible, but it does mean that frameworks such as IDT are of little use when discussing most of the existing lie-detection research. In the next chapter, the verbal and physiological clues to deceit are discussed with reference to this research. A further examination of IDT and naturalistic deception is deferred until Chapter 6, where theories of CMC, deceit, and lie-detection are brought together and IDT is explored once again in order to provide some enlightenment towards uncovering the central mechanisms that drive computer-mediated deceptive encounters.

## Chapter 3

# Detecting Deceit Through the Analysis of Verbal and Physiological Clues

---

Lie-detection research can analyse diagnostic clues at a verbal, nonverbal, or physiological level of analysis. Of these channels, the verbal and physiological research is probably of the most use to the domain of text-based CMC (where nonverbal channels are elided<sup>1</sup>). Verbal behaviour can be detected through the written word and physiological clues may be exhibited through low-level behavioural changes at the computer interface. A discussion of the physiological clues to deceit will be limited to a review of two of the most popular polygraph tests (the Guilty-Knowledge Test and the Control-Questions Test). These will be examined at the end of the chapter. The first part of this chapter concerns the verbal approach.

It has been suggested, that it is too difficult to detect lies through the verbal channel because of the ease at which a person can control their spoken behaviour (Ekman, 2001). After all, a person can control what they say far more easily than they can control certain automatic nonverbal behaviours. It is perhaps this reason why so much of the lie-detection literature focuses solely on the underlying nonverbal behaviours that correlate with deceit.

The significance of verbal control is certainly relevant when a person is required to provide a very short deceptive response, but it is less important during longer periods of interaction. Social perception research has been criticised for its focus on cognitive mechanisms whilst underplaying the importance of the actual perceptual material on which these mechanisms work (Zebrowitz, 1990). This distinction is fundamental to verbal lie-detection research whereby the important difference between the presentment of a truthful or fabricated statement, lies not in the mechanisms at work but in the details related. The difference between encoding and remembering a truth or a lie does not hinge upon the practically identical encoding and retrieval functions of memory, but on the perceptual details encoded in the first place (Johnson and Raye, 1981). In short, the question for verbal lie-detection research is not whether a person can control what they say, but whether or not they know what it is they are supposed to control. In fact, a deceiver's ignorance of the true verbal clues to veracity, and an independent observer's contrasting knowledge of these clues is the precondition for any effective verbal-lie detection technique. The reason why the verbal channel is such a promising route for lie-detection is that the verbal characteristics of truthful statements are largely unknown by people (Ekman, 2001)

---

<sup>1</sup> There is a form of 'textual paralanguage' that exists within CMC, although no research has yet examined the relationship between this 'nonverbal channel' and the performance of deceit.

and that these characteristics may be too numerous for an average deceiver to manipulate them all simultaneously. A selection of techniques is reviewed within this chapter that have attempted to identify the verbal clues to deceit and veracity.

In the early literature, researchers investigated the relationship between verbal characteristics and the behavioural correlates of deception. Vrij (2000) identifies seven significant verbal characteristics that are present within this literature. *Negative Statements* may follow from negative emotions. For example, anxiety at the fear of being detected (detection apprehension) may cause a deceiver to respond with negative utterances such as denials or disparaging statements. The complexity involved in lying may cause an inexperienced liar to make *implausible* statements. The liar may provide excess *irrelevant* information in order to substitute for an abrupt answer, or to draw attention away from the item being lied about. Guilt can cause deceivers to try to disassociate themselves from their lies by either *over-generalising* or by avoiding the use of *self-referencing* words. *Direct answers* and *short responses* may also be symptomatic of the complexity of lying by economically restricting the amount of deceptive information that needs to be managed.

Vrij (2000) provides a literature review of eighteen studies pre-1990 that examine these characteristics. The studies show a relatively high consistency suggesting that there could be a clear relationship between verbal behaviours and deception. At the moment, the foremost technique for verbal lie-detection is Statement Validity Assessment (SVA), a technique that uses a variety of verbal clues that are more sophisticated than those surveyed by Vrij (2000). SVA will be discussed in this chapter as well as Reality Monitoring (another content-based approach). Towards the end of the chapter, two less-common systems for the inspection of verbal clues will be presented including the commercial technique SCAN and the simple assessment of lexical diversity.

### **3.1 Statement Validity Assessment**

Statement Validity Assessment is a three-part process that assesses the veracity of eyewitness statements within a forensic setting. In the sections that follow, the history of SVA is reviewed along with a description of the techniques and methods used within the SVA framework. The discussion then proceeds to methods used to research and evaluate SVA.

#### **3.1.1 History of SVA**

Statement Validity Assessment is a technique used to assess the veridicality of eyewitness accounts. It was set forth as a refinement to its predecessor, Statement Reality Analysis (SRA) (Stellar, 1989). The technique's seminal proponent Udo Undeutsch has provided an account of the origins of SRA and this treatise serves as the foundation for this section (Undeutsch, 1989).

SRA was developed to provide a more objective means by which a witness' statement could be assessed for its 'truthfulness'. Originally, the technique was intended to be used in the

assessment of a child's testimony in cases of sexual abuse. Stern (1902) found significant individual differences in the ability to report the details of an event in retrospect, and he found that perfect recall was rarely attained. Undeutsch highlighted that in cases of sexual abuse, the child's statement is often the pivotal piece of evidence that determines the outcome of the allegation. Early techniques for assessing the veridicality of statements were often ineffective because 'reasonable doubt' could usually be inferred about the witness' memory. Psychological assessments would often base findings upon early dogma concerning women and teenagers. The general inability of women to recollect accounts accurately was often cited, along with the supposed adolescent disposition whereby teenagers were liable to confuse reality with fantasy as a product of heightened sexual cognisance. Undeutsch (1989) reports that 90% of psychological assessments between 1900 and 1930 concluded reasonable doubt, often contrary to the experience of the court.

Undeutsch warns of misguidedly assessing the credibility of a statement solely on the basis of a credible disposition. He suggests that a person with a good reputation may be strongly motivated to lie in order to protect that reputation, and warns against evaluating the veracity of a person as invariant. Clearly there was a need to develop a more objective measure of statement credibility.

First of all, it is important to define credibility. Undeutsch defines credibility as having two dimensions, the cognitive (the ability to recount) and the motivational (the willingness to recount). Undeutsch (1989) states that the motivational element "is a pivotal one in virtually every criminal proceeding" (p.107). Therefore, any technique that assesses the credibility of a statement should focus upon this motivational element and should not hinge upon the remembrance of detail.

In 1954, the West German Supreme Court assessed the utility of psychological assessments in the determination of child sexual abuse cases (Vrij, 2000). Undeutsch appeared at the hearing and subsequently a ruling was passed in 1955 that instructed investigators to carry out psychological interviews and credibility assessments in nearly all child sexual abuse cases. Vrij (2000) reports Arntzen's 1982 estimate that by that date 40,000 cases were subjected to this analysis.

Undeutsch's Statement Reality Analysis aimed to assess the truthfulness of a statement rather than the disposed 'credibility' of a witness. SRA was composed of a structured interview followed by a content analysis. The content analysis consisted of sixteen categories (criteria) that were indicative of truthful statements and these have been reproduced below.

### ***1. Criteria Derived from Single Statements***

#### ***1. General, Fundamental Criteria***

- a. Anchoring (embodiment) in time and place***
- b. Concreteness (clarity and definiteness, specification, vividness)***
- c. Wealth of reported details***



- d. *Originality (individual depiction, more than clichés, trite and stereotyped phrases)*
- e. *Internal Consistency*
- f. *Mentioning of details, specific of the particular type of sex offence*

2. *Special Manifestations of the Aforementioned Criteria*

- a. *Reference to details that exceed the probable capacity of the witness.*
- b. *Reporting of subjective experiences*
- c. *Mentioning of unexpected complications*
- d. *Spontaneous corrections, specifications, complements*
- e. *Self-disserving interspersions*

3. *Negative or Control Criteria*

- a. *Lack of internal consistency*
- b. *Lack of consistency with the laws of nature or science*
- c. *Lack of external consistency (discrepancy with other incontrovertible facts)*

**II. Criteria Derived From Sequences of Statements**

- a. *Lack of Persistence (stability, steadfastness)*
- b. *Prior inconsistent Statement*

*Adapted from (Undeutsch, 1989)*

The content analysis was empirically grounded upon what was later to be coined the 'Undeutsch Hypothesis', "that a statement derived from a memory of an actual experience differs in content and quality from a statement based on invention or fantasy" (Vrij, 2000; p.116).

Subsequent development of the technique led to the emergence of Statement Validity Assessment. This adaptation by Steller and Köhnken (1989) used nineteen criteria that comprised what is now called the 'Criteria-based Content Analysis' (CBCA). Steller and Köhnken provided a full procedure for assessing veridicality that incorporated an interview, CBCA, and also a set of questions (the Validity Checklist) that evaluated whether or not the procedure would be appropriate for the analysis of a particular statement (Steller, 1989). SVA differed from SRA by drawing more attention to the interview characteristics (such as the potential for suggestibility) as well as other sources of influence (operationalised through the Validity Checklist).

Since its original inception in Germany, SVA has only appeared in the UK and the US over the last fifteen years and its forensic application has been tentative. To date, it is not used in the UK courts and opinion about its utility varies in the US between states (Vrij, 2000). Researchers in these countries were surprised at the limited amount of existing scientific data examining the SVA technique (Vrij, 2000). Vrij (2000) reports the existence of only fifteen studies that examined CBCA comparing genuine and fabricated statements on the basis of the individual criteria. It is perhaps this lack of empirical wealth that impedes SVA's forensic application in this country.

Undeutsch (1989) himself sees little value in conducting laboratory work to test SVA because of the poor ecological validity that would be obtained by the results. He objects to the decisions of the UK and the US courts to not take up SVA in the same way as Germany had done. He speaks with reference to the ease with which a defense attorney can utilise (non-SRA) psychological evidence to cast doubt on a statement's validity. The reservations shown by the UK and the US mirrored those of Germany, and that it was only a matter of time until those countries adopted a standardised psychological assessment technique as well.

Since Undeutsch made this remark, there have been a number of studies that suggest that SVA (at least in its current form) should be used with caution. These studies and their reservations will be discussed in a later section. Suffice it to say that the apparent success of the technique in Germany is not a sufficient measure of its internal validity. As will be discussed later in this chapter, the technique is prone to a truth bias that may simply correlate with the hitherto unknown 'ground truth' of the veracity of child abuse cases. Nonetheless, the technique still has a strong support from the laboratory studies that do exist, especially when the relative success of other 'truth validation' or 'lie-detection' techniques are considered.

In summary, SVA was originally developed out of SRA as a technique for assessing the veracity of statements from alleged victims of child abuse. The technique was popularised in West Germany but has been slow on the uptake in the UK and the US. It seems as though much more empirical work still needs to be undertaken in order to assess the validity of the technique. The technique is based upon the Undeutsch hypothesis that postulates qualitative and quantitative differences between experienced and fabricated accounts. The SVA procedure is comprised of three parts, the interview, the CBCA, and a Validity Checklist and these will be discussed further in the sections that follow.

### **3.1.2 Overview of Methodology**

The SVA procedure begins with a carefully conducted interview. The transcript of this interview then needs to be consistently and reliably coded using CBCA. Raskin and Esplin (1991a) suggest "a structured interview designed to maximize the amount of information provided by the child witness and to minimize contamination of its contents by the interviewer" (p.273). A good way for the interviewer to start along these lines is by asking the eyewitness to describe their story from start to finish without interruption. This will hopefully elicit a free narrative that is uncontaminated by the interviewer. The structured interview must be conducted by a trained interviewer and should take between twenty and forty-five minutes. An open-ended style of questioning should be used that avoids leading the witness. The interviewer would need to be familiar with the case itself, and this preparation is a *sine qua non* of a successful interview as it would be almost impossible to conduct it otherwise.

Central to the Raskin and Esplin interview technique involves the investigation of a set of alternative hypotheses that reflect other explanations for the allegation. The authors provide some examples of alternative hypotheses including alternative perpetrators, outside influence,

revenge, or fantasy. In order to evaluate these hypotheses, the interview should be scheduled such that this sort of information can be obtained in advance.

Raskin and Esplin (1991a) present four types of question that can be used during the interview, cue, probe, direct, and suggestive. Cue questions are open-ended questions that refer to something specific in the account. Probes are used to test the alternative hypotheses whereas direct questions aim to elicit a specific piece of information without suggesting a possible answer. Finally, suggestive questioning can be used at the end of the interview in order to test the overall suggestibility of the witness. This helps the investigator to rate the overall validity of the procedure.

Much of the empirical work that surrounds SVA does not yet utilise a structured SVA interviewing procedure such as the one presented by Raskin and Esplin. Over the last few years, there has been an elevated interest in the cognitive interviewing technique and its role in deception detection. This will be described further in section 3.1.4.2.

The Criteria-based Content Analysis is at the heart of the SVA technique. In its present incarnation, it consists of nineteen criteria that are used to rate a transcript. The full list of criteria was compiled by Steller and Köhnken (1989), but a number of other authors have used either a subset or a superset of a subset of these. Each criterion is evaluated on some form of discrete ordinal scale, usually comprising of three levels such as 'Absent', 'Present', or 'Strongly Present' (Steller, 1989). The final score usually represents some form of weighted or un-weighted totaling of all the criteria. Each criterion represents a typical characteristic of a genuine statement but an individual statement may or may not contain certain criteria. Certainly, the absence of any one criterion does not indicate fabrication as the technique is strictly concerned with *truth* validation.

What follows is a description of each of the criteria. These descriptions are adapted from those provided by Vrij and have been included here primarily as a convenience for the reader, (see Vrij, 2000; pp. 116 to 122 for the original descriptions).

### ***I. General Characteristics***

*The first three criteria refer to the statement in its entirety.*

#### ***1. Logical Structure***

*The coherence and logical consistency of the statement. Genuine statements tend to lack inconsistencies, contradictions, or other discrepancies.*

#### ***2. Unstructured Production***

*Genuine statements are typically not characterised by a linear temporal structure. Vrij reports that unstructured accounts are common when people are distressed, but the effectiveness of the criterion can be mediated if the story has been told a number of times before.*

3. *Quantity of Details*

*Genuine statements tend to incorporate a variety of details. This criterion assesses the richness of detail in the whole statement.*

**II. Specific Contents**

*The following criteria pertain to certain passages in the statement that are supposed to indicate 'concreteness' and 'vividness'.*

4. *Contextual Embedding*

*This criterion is present if the statement contains temporal and spatial markers and if actions are integrated into daily activities (i.e. situated in context).*

5. *Descriptions of Interactions*

*This criterion is present if there exists interactive elements between the witness and the perpetrator.*

6. *Reproduction of Speech*

*Genuine statements also contain instances of utterances that occurred during the alleged event. The criterion is only fulfilled if the utterances are in their original state and if the speakers can be clearly inferred.*

7. *Unexpected Complications During the Incident*

*This criterion is present if any complications are reported in the statement.*

8. *Unusual Details*

*This criterion denotes the existence of any unusual details contained in the statement.*

9. *Superfluous Details*

*This criterion rates the existence of details that are connected to the event but do not add weight to the allegation.*

10. *Accurately Reports Details Misunderstood*

*This criterion probably has a greater relevance for children than adults. It is fulfilled if details are reported that indicate an action that lies outside the understanding of the witness. This could include a young child describing a sexual act, or possibly a witness describing a technical procedure where the function of that procedure is either unknown or misperceived.*

11. *Related External Associations*

*This criterion is fulfilled if the statement contains information, which, although not a specific detail of the event, is relevant to that event.*

12. *Accounts of Subjective Mental State*

*The inclusion of feelings and thoughts or cognitions at the time of the incident.*

13. *Attribution of perpetrators Mental State*

*The inclusion of perceptions concerning the perpetrators feelings, thoughts and cognitions. This criterion is again dependent on there being a 'perpetrator' in the transcribed event.*

**III. Motivation-related Contents**

*The motivation-related contents are indicators of how the statement was constructed with regard to the goal of being truthful as opposed to being deceptive. They are related to the principle that a deceiver may wish to suppress those misperceived stereotypical behaviours associated with lying.*

14. *Spontaneous Corrections*

*This criterion is present if corrections or supplemental information are provided without any intervention by the interviewer.*

15. *Admitting Lack of Memory*

*Admissions of lack of memory, offered supplementally to an open question, fulfil this criterion. Simple, flat admissions, in response to closed questions are not included.*

16. *Raising Doubts About Ones Own Testimony*

*This criterion is present if concern is shown that a part of the statement may be incorrect or unbelievable.*

17. *Self-deprecation*

*This criterion is present if details are mentioned that are personally unfavourable, or that imply some fault on the part of the witness.*

18. *Pardoning the Perpetrator*

*Present if excuses or other details are provided that mediate the onus of guilt away from the perpetrator. It is reasonable to assume that a deceiver interested in false incrimination would not desire to mediate, but to militate such a perception.*

**IV. Offence Specific Elements**

19. *Details Characteristic of the Offence*

*This criterion is present if the statement describes a sequence of events that follow a format typical to that particular crime. Such a judgement would need to be made by experienced professionals who are aware of the typical modus operandi of certain types of crime.*

A popular set of guidelines for applying the content analysis was put forward by Steller (1989). Steller and his colleagues produced detailed descriptions and working sheets and decided upon the following three rules.

1. *That any repetition in the statement does not increase the rating of a relevant criterion.*
2. *Any passage in the statement can fulfil more than one criterion.*
3. *Only content that is relevant shall be considered.*

Despite such efforts to standardise CBCA, the technique still requires the raters to apply a certain amount of subjective judgment. Official criterion definitions, rules, and procedures have helped to standardise the technique, but these procedures have yet to reach maturity. The CBCA provides a means of assessing the veridicality of a statement, but as the CBCA is not a standardised test (Vrij, 2000), there are a number of factors that govern how the CBCA score should be interpreted. Specifically, CBCA could only be considered a valid test under a certain set of circumstances. Steller (1989) comprised an eleven-item list that defines those circumstances where a CBCA analysis is inappropriate. This is known as the Validity Checklist (VCL) and it comprises the third component of the SVA. The Validity Checklist has been adapted here for convenience (the reader is referred to Steller, 1989 for the original definitions).

#### ***I. Psychological Characteristics***

1. *Witness used inappropriate language or demonstrated inappropriate knowledge that was beyond their capacity.*
2. *Witness exhibited inappropriate affect with regard to the statement contents.*
3. *Witness seemed susceptible to suggestion.*

#### ***II. Interview Characteristics***

4. *Questions posed during the interview were suggestive, leading or coercive.*
5. *Interview was inadequate in other aspects.*

#### ***III. Motivation***

6. *There exists a questionable motive behind the witness coming to testify.*
7. *The original disclosure or testimony was attained under questionable conditions.*
8. *There exists outside pressures to deceive.*

#### ***IV. Investigative Questions***

9. *The statement seems to run contrary to the 'laws of nature'.*
10. *The statement seems inconsistent with other statements.*
11. *The statement seems inconsistent with other evidence.*

If a particular case exhibits many of the items mentioned in the VCL, then the use of CBCA is probably inappropriate. For example, the characteristic described in item one could indicate that the witness has been tutored in a particular story by a third party. Even if the story itself is true (and it is important to reiterate that the VCL does not assess veracity), it will still be the case that a tutored eyewitness will not be providing the kind of naturally occurring statement upon which the empirically grounded CBCA technique depends.

### 3.1.3 Overview of Empirical Studies

There has been a dearth of empirical studies that examine SVA and its flagship CBCA. Scientific studies that examine the procedure can be conducted in the field or in the laboratory. As would be expected, the perennial critiques of each of these methods can be debated at length. Field studies have been criticised for their inability to ensure the accuracy of ground-truth and laboratory studies have been criticised for their insufficient level of ecological validity. A fuller discussion of the relevant research issues pertinent to these studies has been deferred until the next section. In this section, some of the existing empirical work has been reviewed with regard to the overall success at detecting genuine and fabricated accounts. The purpose of this section is to describe a few of the principal field and laboratory studies that illustrate the typical modus operandi of SVA research.

#### 3.1.3.1 *The Experimental Approach*

Experiments that assess SVA or CBCA typically involve two parts; the generation of stimulus material and the codification of that material by trained CBCA raters. The stimulus material usually comprises videotapes or interview transcripts where the participant has been asked to tell the truth on one occasion (the control condition) and to lie on another. The power of the experimental approach is the ability to record the 'ground truth' (i.e. which story was the true one). Often the participant is asked to perform some sort of mock crime beforehand and this serves as the interview topic. The second part of the study involves the CBCA codification of the transcripts by a number of trained raters. As with many other experiments of this kind, university undergraduates are often used as participants. These experiments have typically reported the discriminatory power of CBCA, along with the prevalence of certain criteria over others.

In 1992, Landry and Brigham published an experiment that demonstrated the effectiveness of CBCA training on the ability to detect deception (Landry and Brigham, 1992). The stimulus material was drawn from a pool of personal incidents related by participants. Each participant was filmed giving short descriptions of either experienced or fabricated events that were in some way traumatic, involved and that were characterised by feelings of loss of control. Half the videotapes were true accounts and half were fabrications that were written by the investigators. Sixty-four undergraduate raters were allocated to one of four conditions following a 2 (Training:No Training) X 2 (Videotape:Transcript) experimental design. The trained raters used only fourteen of the CBCA criteria as some were deemed irrelevant in this scenario (c10, c18, and c19)<sup>2</sup> and others were found to be too difficult to apply (c2 and c11). The results showed that trained raters performed significantly better overall (55.3%) than untrained raters (46.9%) but only slightly better than chance (50%). Nonetheless, this low overall CBCA accuracy rating represents a bias to truth detection such that the trained raters judged 75.4% of true stories correctly as opposed to 58.6% for the untrained raters. Surprisingly the accuracy rates were

---

<sup>2</sup> For brevity, the individual CBCA criteria will be referred to by their number.

also higher for videotape ratings as opposed to transcript ratings. The authors suggest that the improved results may be the result of the raters using various nonverbal cues to help them with their judgments. However, this conclusion should be evaluated with caution considering the body of evidence on the falsely perceived nonverbal clues to deception (see Ekman, 2001).

Zaparniuk, Yuille and Taylor (1995) published a laboratory study that, like the Landry and Brigham study three years earlier, was concerned with the application of CBCA to the statements of adults. To produce the stimulus material, forty undergraduates in a 2 (Truth: Fabrication) X 2 (Videotape:Verbal description) between-subjects design were used. The participants either were shown a videotape or were read a verbal description of a two-minute scenario. The participants in the fabrication condition undertook a stepwise interview whereby they were told to relate the story as though they had actually experienced the event. The CBCA rating was undertaken by two trained and one expert rater who were using two decision rules to supplement the codification procedure.

- **Decision Rule 1:** Identify the presence of c1, c2, c3, c5, plus any remaining two criteria.
- **Decision Rule 2:** Identify the presence of c1, c2, c3, plus any remaining four criteria.

The purpose of the decision rules was to assess whether or not they could be more effective at discriminating true and fabricated statements than simple criteria tallies. However, the danger of this type of strategy is that the use of specific decision rules could create a discriminatory bias towards a particular type of scenario. This is especially applicable in the case of the Zaparniuk et al. study, where a rather subtle form of deceit was being used in the fabrication condition. Interestingly, after codification c2 by itself had an 80% accuracy rate, and the two decision rule variants yielded 78% and 73% accuracy rates respectively. However, c2 also had a relatively lower inter-rater reliability and the authors suggest that the decision rule approach will probably be more reliable than using c2 by itself.

Ruby and Brigham (1998) conducted the first study that examined ethnicity and its affect on CBCA scores. The stimulus material for the study comprised transcripts of six White (Eurocentric condition) and six Black (Afrocentric condition) participants giving statements of two personal events, one true and one fabricated. As with other studies, the participants were informed that the events should be characterised by a loss of control. The participants were given one or two days preparation time to rehearse their accounts. The study found that CBCA training did not enhance in the truth detection of White speakers over Black speakers, but five of the criteria (c3, c6, c12, c14, and c17) did in fact appear more in the statements of Black speakers than White speakers. The study also identified a strong truth bias in CBCA that led to a high amount of false-positive errors. A subsequent discriminant analysis was able to reduce such false-positives, but only at the expense of an increase in false-negatives. Like many other investigators, Ruby and Brigham concluded that it was too premature at that stage to use CBCA as a diagnostic tool to detect veracity.



### **3.1.3.2 The Field Study Approach**

In comparison to the experimental approach, even less field studies have been conducted that involve SVA. This could be because of the difficulty in verifying the ground truth of each case under examination. However, the primary advantage of field analysis is ecological validity. The results from experimental studies are often criticised on the grounds that the participants were not sufficiently motivated to lie well, or that the subject matter did not have a sufficient degree of personal, contextual relevance. Field studies are usually conducted post-hoc using videotapes from actual police interviews and are coded in the same way as for experimental studies.

Anson, Golding, and Gully conducted a study that used stimulus materials from actual child abuse cases in Utah (Anson et al., 1993). Twenty-three statements (with videotapes) were selected from a pool of 512 cases that were investigated between the years 1986 and 1989. In this instance, the two expert CBCA raters used the videotapes as stimulus material as opposed to the transcripts. The main purpose of the study was to assess inter-rater reliability between the two CBCA experts. The authors found that only nine of the criteria were scored with accurate reliability, and only four of these criteria were found to be present less than 10% or more than 90% of the time (suggesting that they had been clearly understood by the raters). The authors provide a number of reasons for this disappointing result, mainly focusing around training issues (such as the inadequacy of the criteria definitions) and the interview conditions that were more an artifact of the research method than of SVA itself.

Another CBCA study conducted in a field setting was undertaken by Lamb and his colleagues (1997). The authors cite two previous field studies undertaken in the US, that of Boychuk, and Raskin and Esplin (Boychuk, 1991; Raskin and Esplin, 1991b). The authors address the critique that these studies had struggled to establish the ground truth by some objective, independent measure by which to compare against CBCA scores. This study, conducted in Israel in the Hebrew language, used a multifaceted technique developed by the authors in order to determine the ground truth of the sample material. Ninety-eight cases were taken from a pool of 1187 interviews. The CBCA ratings were conducted from transcripts by two native Hebrew speakers and utilised only fourteen of the criteria. The focus of the study was on validity rather than reliability, but raters were still trained to 90% inter-rater reliability levels. The authors found significant associations between case plausibility and CBCA ratings as well as a significant correlation between CBCA scores and victim age. Nonetheless the results were less dramatic than those presented by Boychuk (1991) and Raskin and Esplin (1991b). The authors suggest that this difference results from a greater level of ecological validity, but they also acknowledge that the interviews were not conducted in a manner conducive to SVA (fewer open ended questions). The results also show significant heterogeneity in the classification power of the criteria (c1 was consistently rated as high across all of the conditions). Based on their results the authors conclude that overall, "the level of precision clearly remains too poor to permit the designation of CBCA as a reliable and valid test suitable for use in the court room" (p.262).

### **3.1.4 Interviewing for SVA: Research Issues**

The transcript used for the CBCA is elicited by means of a structured interview. Interviews are subject to a variety of influences that dictate exactly what information is revealed. It will be recalled that the Undeutsch Hypothesis posits both quantitative and qualitative differences between genuine and fabricated statements. Therefore, it is important that the SVA interview can provide a source of information that does not confound these differences through artifacts of the procedure. For example, hypnotic interviews have been criticised because they induce a highly suggestive affective state. This state increases the likelihood that the statement will be corrupted by the interviewer (Putnam, 1979). An interview transcript is simply an interpretation of the events as remembered by the interviewee. It is essential that this interpretation does not create a bias that will lead either to false positives or to false negatives. Similarly, it is also essential that the interviewing process provides sufficient data for the CBCA to take place. This section reviews the findings that have been revealed using the investigative interviewing technique with respect to those issues pertinent to SVA.

#### ***3.1.4.1 Principles Underlying the SVA Interview***

The SVA interview should be designed to minimise data corruption and to maximise both data quantity and quality. The style of questioning adopted and the interpersonal relationship between the interviewer and the eyewitness is central to these tenets.

The question that is posed here, is what is the ideal interviewer-respondent relationship to maintain for SVA? As SVA has both forensic and empirical roots, the answer to this question may lie in the techniques that are employed by police to interview eyewitnesses. The clichés contained in old detective stories, that involve the suspect sweating under lamplight whilst simultaneously cowering in front of an aggressive interrogator, illustrate the worst violations of the validity checklist. The principle behind this type of interrogation is that if a suspect is guilty of a crime, then the weight of the whole procedure will eventually manage to 'break down' the façade of innocence. In Orwell's 1984 dystopia, the suspect through a similarly strenuous psychological procedure provides a false confession, and then finally ends-up believing in that confession (Orwell, 1984). To what extent do these two scenarios echo similar characteristics found in police interviews?

It is fair to say that in the UK (and indeed many other nations) the process of interrogation no longer relies upon extracting a confession but upon obtaining evidence (Milne and Bull, 1999). The reasons for this are partially to do with the danger of attaining Orwellian false confessions, and partially because this sort of confrontational approach runs contrary to the goal of maintaining an open relationship (Milne and Bull, 1999). Evidence against the confession-driven approach can also be found empirically. Baldwin (1993), in an analysis of six hundred audiotaped interviews found that in only nine cases was there a reversal of stance by the interviewee that was attributable to the persuasive skills of the interviewer.

One problem of putting a witness under pressure is that it is likely to cause the witness to withdraw. This will in turn restrict the amount of information that can be made available for the CBCA. Similarly, as with modern day police interviews, the purpose of the SVA interview is to obtain information and not a confession and therefore it is not necessary to employ the kind of trickery that was suggested in the early literature (for example in the book produced by Inbau, Reid and Buckley, 1986). It would seem that in order to extract as much information as possible from the eyewitness, the interviewer should maintain a non-accusatory relationship of trust. This may be helped by clarifying the purpose of the interview, and the credentials and identity of the interviewer from the outset (Milne and Bull, 1999). This might appear to have the negative effect of relaxing a deceiver and making them less likely to make mistakes, but the influence of this is limited because SVA is not affect-driven.

In establishing this relationship, it is necessary to consider the power distribution between the interviewer and interviewee. A traditional structured interview requires the interviewer to drive the discourse along a predetermined schedule. The problem with this strategy is that the interviewee may be being forced to tell their account in a manner that does not reflect the way in which the information was encoded into memory and this is not conducive to good recollection (Tulving and Craik, 2000). It should therefore be suggested to the eyewitness that it is their narrative and not the interviewer's questions that will determine the structure of the interview.

An effective relationship can be maintained through the establishment and management of rapport. This can be achieved by using nonverbal skills and active listening. The interviewer must allow a sufficient amount of interpersonal space (proxemics) between them and the interviewee and should adopt an appropriate posture and paralinguistic style (Milne and Bull, 1999). Milne and Bull also suggest that when framing their questions, the interviewer should use language that is similar to the words and expressions used by the interviewee. Experimental research has highlighted a halo effect that surrounds the 'general nonverbal expressiveness factor' (Riggio and Friedman, 1986; Friedman et al., 1988). This research has found that people who have a high level of nonverbal expressiveness are usually attributed more positive traits than are their less expressive equals. Highly expressive professionals are typically rated as being more competent and this can help to promote the relationship between the interviewer and interviewee. The flipside to high expressiveness lies in the observation that highly expressive people can have difficulties in showing neutrality. This would naturally have implications for a SVA interview where the interviewer must be careful not to convey their own opinions regarding the truthfulness of the account.

The questioning techniques that promote maximum memory retrieval are generally those that are open-ended. Fisher, Geiselman and Raymond conducted a study that involved the analysis of a number of recorded police interviews (Fisher et al., 1987). These interviews usually begin with an open-ended request for a description of "what happened?". On average, the interviewee was interrupted after only seven-and-a-half seconds. The authors suggest that these interruptions might cause the interviewee to infer that their responses are too lengthy and that

they must shorten their answers to future questions (as a form of conversational regulation). As uninterrupted free responses are usually rich in data, it would seem prudent to avoid such disturbances as much as possible.

Aside from the free narrative, the use of open *and* closed questions can still be appropriate for investigative interviewing (Milne and Bull, 1999). Open questions can be used to probe for more information surrounding a feature contained in the free narrative. Closed questions can be used to obtain facts concerning particular items in the account, or to provide control over the direction of the interview. Milne and Bull (1999) state a preference for open ended questions, as the use of short-answer questions “causes the interviewees to be passive, decreases their concentration and therefore can result in less recall” (p.23).

#### **3.1.4.2 The Cognitive Interview**

The cognitive interview (CI) is a technique that incorporates many of the good practices outlined in previous sections. The CI is a marriage between cognitive and social psychology, and between maximising memory retrieval and managing social communication (Memon and Higham, 1999).

CI was created by Fisher and Geiselman (1992) in response to “a pressing need for an investigative interviewing technique which was based on established psychological research” (Milne and Bull, 1999; p.33). Since its original incarnation the authors have produced an enhanced version of the technique (ECI) that provides a template for an interviewing schedule. This schedule has been reproduced below.

##### **1. Introduction**

*Control eyewitness anxiety, develop rapport, tell eyewitness to actively volunteer information and not passively wait for interviewer to ask questions. Explicitly request detailed information. Tell eyewitness not to edit her thoughts. Tell eyewitness not to fabricate or make up answers. Convey that the interviewer expects the eyewitness to concentrate intensely.*

##### **2. Open-ended Narration**

*Recreate the general context. Request narrative description and do not interrupt. Provide a long pause after eyewitness stops speaking before asking the next question. Identify eyewitnesses images and develop a tentative probing strategy.*

##### **3. Probing Memory Codes**

*Re-emphasize the importance of eyewitness concentration. Recreate the context of a specific event and ask the eyewitness to close her eyes. Ask an open-ended, framed question. Request detailed description and do not interrupt eyewitnesses narration. Take detailed notes. Provide a long pause*

*after eyewitness stops speaking before asking follow-up questions. Exhaust image for information not included in narration. Probe remaining images and re-probe images activated earlier.*

**4. Review**

*Review for eyewitness from interviewer's memory or notes. Speak slowly and deliberately. Ask eyewitness to interrupt immediately if she remembers new information or if there are errors in interviewer's review. If any new leads develop then probe the relevant information.*

**5. Close**

*Collect background information. Remind eyewitness to call when she thinks of new information. Create a positive last impression.*

The Cognitive Interview is comprised of four primary techniques. First, the interviewee is instructed to report all details of the event in question, no matter how incidental or unimportant. Secondly, the interviewee is asked to try to recreate (but not fabricate) the event in their minds. By doing this, the interviewee is accessing the remembered physical and mental information of the event and subsequently activating any potential retrieval cues. After a short period, the interviewer then attempts to elicit the free narrative. During this period, the interviewer will be taking notes that will form the basis for later questioning. The third stage involves 'probing memory codes' which involves asking the eyewitness to recall and visualise the key 'images' that were revealed during the narration. This will be followed up with an open question that is intended to explore the image in greater depth. These questions can be supplemented with varied retrieval methods that aim to promote recollection. Such methods include asking the witness to relate events in a different temporal order or from a different perspective. The whole process concludes with a review of the eyewitness' statement before the interview is terminated.

Some investigators have explicitly examined the benefits of using the cognitive interview as a part of SVA. Steller and Wellershaus (1996) argued that as the cognitive interview enhanced memory it would subsequently affect the content and structure of eyewitness accounts. This would then confound CBCA, as the technique was derived empirically from statements that were elicited by other means. Köhnken and his colleagues (1995) believed that the effect of the CI would be to enhance the quality of both truthful and fabricated statements, but that it would leave CBCA's discriminatory power intact. They devised a typical SVA-style experiment involving six participants that had no previous interviewing experience. The interviewers were either trained in CI (two, two-hour sessions) or in another structured interviewing methodology and fifty-nine interviews contributed to the analysis. The authors found that there were significantly more correct details identified in the CI condition, and no significant differences in incorrect details or accuracy across conditions. Nonetheless, they did find that under the CI, interviewees did tend to confabulate more. The interviews in the CI condition yielded

approximately 35% more correct information than those in the structured condition. With respect to CBCA, five criteria were found to rate significantly higher in the CI condition (c8, c9, c10, c15, and c16). Even so, the authors found that the CI based analyses were still able to discriminate successfully between truthful and fabricated accounts.

### **3.1.5 CBCA and the Validation Checklist: Research Issues**

The previous section examined the SVA interview in greater depth. In this section, the other two elements of SVA will be addressed, the CBCA and the Validity Checklist. Researchers face a number of challenges when undertaking experimental or field research concerning SVA. Those challenges, relating to validity, reliability, and realism have been discussed below.

#### **3.1.5.1 Theoretical Support**

Theoretical reflections on CBCA go little beyond the Undeutsch hypothesis in positing qualitative differences between truthful and fabricated statements. This hypothesis has been elaborated into the assumption that somebody who has experienced an event is more likely to remember a broad range of criteria-relevant information. The most theoretically grounded criteria are the motivation-related criteria and these were developed on the basis that a truth-teller would not have the same motivation to appear veracious than the deceiver, and so is less likely to suppress those verbal behaviours commonly believed to indicate deceit or discomfort (Ruby and Brigham, 1997). The difficulty of an empirically driven research approach is that it makes it difficult to derive new criteria, or even to identify what are the most pertinent facets of the existing criteria. This has led some researchers to attempt to consolidate CBCA with existing psychological approaches based upon episodic memory (c.f. Sporer, 1997).

#### **3.1.5.2 Training**

All studies that assess SVA must carefully consider the role and effect that rater training has upon CBCA scores. It is difficult to support claims about the reliability of the method if raters have not been adequately trained for the task. When considering this issue, a number of questions need to be raised including:

- *What is the minimum amount of training required for CBCA?*
- *At what point does continued training make no significant difference to CBCA scores?*
- *How can training be standardised in the research community?*
- *How is training evaluated as being sufficient?*

Vrij (2000) reports that the German Psychology and Law Society are currently considering a means for formalising a SVA training programme but at the moment no such programme exists.

Training typically involves some form of explanation of the methodology, a practice rating session and some form of feedback or comparison process with a trained CBCA expert.

Training is usually deemed sufficient once a certain level of consistency has been obtained between the trainees and the expert.

Due to the indeterminable effects of bias, researchers themselves are unable to perform the CBCA ratings in their studies (Steller, 1989 criticises a field study by Michaelis-Arntzen, 1987 on precisely this point). This can mean that raters need to be trained from scratch and often the length and the form of the training session can be strongly influenced by practical concerns (such as the number of raters being trained). The extent to which such concerns are relevant seems unclear. Ruby and Brigham (1998) used a single forty-five minute training session for their 143 raters. They provided their raters with a list of criterion definitions and explanations and tested them for accuracy. Those raters who had correlated with an expert's assessment to a degree of 80% or more were included in the study. This relatively brief training session yielded accuracy results of 68%. In contrast, Sporer (1997) used only two raters. These raters participated in a seminar about CBCA and read a broad selection of literature on credibility assessment. The raters were also given a surprise memory test of the criteria (this has also been used elsewhere c.f. Landry and Brigham, 1992) and were made to rate and discuss forty practice stories. Sporer achieved similar accuracy levels of about 65%.

Adequate criterion descriptions that include examples of the criteria and how they should be applied are an essential training material. A number of studies have yielded results which the authors suspect could have occurred due to confusion about how to apply certain criteria. For example, Landry and Brigham (1992) found in a pilot study that raters were unable to understand or apply consistently the 'Unstructured Production' and the 'Related External Associations' criteria. Consequently, they omitted these two criteria from their analysis. In fact, this may not have been entirely due to the descriptions themselves but also due to the short training time (a forty-five minute session).

A training schedule that has been relatively popular is one that has been used by Köhnken et al. (1995) and Akehurst et al. (2001). In this system, the rater receives a booklet with detailed descriptions of the criteria along with examples. In a first training session, the trainee discusses the descriptions with an expert rater. In the second session, the trainee is taught to rate the criteria on a seven-point Likert scale. The training session is deemed complete when the trainee and the expert differ by no more than one point on the scale. This schedule seems promising as a potential technique to be used in future studies. However, this methodology will depend upon a high inter-rater agreement between experts if the technique is to become standardised. Ideally, a standard set of transcripts should be produced with detailed ratings that have been discussed and agreed upon by all the leading experts in the research community. This sort of manual remains unpublished but is no doubt the goal of the German Law and Psychological societies.

### **3.1.5.3 Establishing the Ground Truth**

CBCA is intended to be a measure of the veracity of a statement. In order to configure CBCA correctly, the objective ground truth is needed as an independent point of reference. Steller (1989) reports that in field validation research, establishing the ground truth against an independent measure is difficult. After all, if a perfect means existed then there would be little need for CBCA. Steller cites a field study by Michaelis-Arntzen (1987) as typifying this sort of problem, where there was insufficient "clarification of the independent criteria used for classifying the testimonies into groups of 'actually truthful' and 'proven untruthful' statements" (p.149).

Raskin and Esplin (1991c; cited in Lamb et al., 1997) used polygraph examinations, medical evidence, eyewitness accounts and confessions in order to distinguish 'doubtful' and 'confirmed cases'. However, according to Lamb et al., not all of the doubtful cases "were supported by independent evidence that the allegation was false" (p.257).

Lamb et al. (1997) subsequently developed "an elaborate multifaceted procedure" for establishing ground truth in their field study. This procedure involved using raters working in parallel to the CBCA raters to establish ground truth via 'Independent Case Fact Scales' (ICFS). The items in the ICFS were (1) medical reports, (2) witness statements, (3) physical or material evidence, (4) suspect statements, and (5) other miscellaneous information. Ratings used all available case evidence and rated these five dimensions on a scale of 'Very Likely', 'Quite Likely', 'Questionable', 'Quite Unlikely', 'Very Unlikely', 'Left No Judgment Possible', 'Was Not Relevant', or that the relevant information was 'Not Obtained'. Naturally, it is difficult to rate the success of the ICFS technique since that would also require another independent measure (and so on recursively). There was a correlation between ICFS and CBCA ratings but such a mutual validation is insufficient.

Parker and Brown (2000) also attempted to classify true and 'unfounded' statements in a controlled manner. Again, other sources were used including medical evidence, admissions of deceit by the alleged, evidence of delusion, and legal corroboration. The authors remarked, "it is, of course, possible that incorrect classifications were made, but every care was taken to ensure accuracy and where there still was doubt cases were eliminated from the analysis" (p.241). Whilst this would seem to be an acceptable strategy, it would still lead to an inevitable sample bias. The authors also gave no indication of what percentage of their classifications might be expected to be incorrect. Unfortunately, it would seem that this sort of information cannot be obtained with any known degree of certainty and this is the underlying difficulty of using field studies to validate CBCA.

Although experimental studies may champion a higher accuracy of ground truth, it cannot be assumed that this information will always be known. Participants may maliciously or inadvertently mix up when they should be telling the truth and when they are supposed to be fabricating. This is less likely to occur in scenarios where participants have to relate a scenario



that is known to the investigators as opposed to a scenario about personal events in which they have no knowledge.

#### **3.1.5.4 The Base Rate of Lying**

It would be easier to assess the general level of ground truth to expect from field studies if the base rate of lying for particular type of allegation were known. The base rate of lying is the percentage of allegations 'in the field' that are actually fabrications. However, an assessment of the base rate of lying is dependent upon measures that are likely to involve some unknown degree of error. As an example, Parker and Brown (2000) observed significant variability in base rates of false allegations of rape. The authors state that such variability is "likely because of lack of a standard conceptualisation of the problem, variability in the quality of investigation and disparate crime classification used in recording rape" (p.238).

Steller (1989) also mentions the importance in understanding the base rate in field studies with relation to the truth bias effect of CBCA. CBCA as a truth validation tool has a tendency to make more false-positive errors than false-negatives. In other words, the technique seems to have a truth bias (Steller, 1989; Vrij, 2000). Subsequently the apparent success of the technique in forensic applications in Germany must be viewed with caution. The success may simply be indicative of a high base rate compounded with a CBCA truth-bias. Additionally as Steller points out "there is no scientific literature discussing the (high) base rate of true allegations in cases of child sexual abuse in connection with the validity of the methods employed [...] Practitioners who would follow the simple rule of assessing every child sex abuse allegation by children as credible would run only a small risk of false-positives [...] by completely avoiding the risk of false-negatives" (p.142).

#### **3.1.5.5 Reliability**

The reliability of CBCA (and of SVA as a whole) depends in part on the factors relating to external validity and training that were mentioned elsewhere in this section. Reliability is usually examined in terms of inter-rater and intra-rater agreement, but the ways in which such evaluations take place, as well as the interpretation of their significance, is dependent on how the process of SVA is conceptualised. This debate has been explicitly addressed in (Horowitz et al., 1997; Tully, 1998; and Horowitz, 1998) and will be examined here.

Horowitz and his colleagues, conceptualise CBCA as a psychometric instrument 'in the making'. In their 1997 article, they argue for the need to increase the attention that is paid to reliability considerations in SVA studies (Horowitz et al., 1997). They explicitly examined inter-rater and intra-rater reliability under a psychometric paradigm. To this effect, three raters were used to evaluate one hundred transcripts and this evaluation occurred twice with no less than three months difference between the two sessions. In these sessions, the criteria were rated on a bipolar scale. They found no significant inter-rater or intra-rater differences in total scores, but further analysis revealed that there was a significant difference in reliability across individual

criteria. Subsequent analysis using a complementary calculation of Cohen's kappa and Maxwell's random error coefficient, showed that some criteria revealed very low, unacceptable reliability scores ('Lack of Memory', 'Superfluous Details', 'Spontaneous Corrections', 'External Associations', and 'Unusual Details') whereas many showed exceptional reliability scores ('Raising doubts', 'Self Deprecation', 'Pardoning', 'Accurate Details Misunderstood', and "Logical Statements"). Overall, they evaluated the performance of CBCA to be satisfactory, but qualify this with a caution that the low scoring criteria should either be removed or redefined if further evidence contains similar reliability scores.

A critique of this means of conceptualising CBCA psychometrically was made by Tully (1998). Tully views SVA as a protocol rather than a psychometric test. Tully suggests that the reliability measures reported by Horowitz et al. were as much reflective of the transcript 'quality' than of the methodology. Horowitz et al. had not graded their sample with regard to the expressive quality of the narrative. Tully suggests that certain criteria may appear to have a poor reliability because a certain 'threshold for criterion clarity' had not been attained by the sample (Tully, 1998). It would seem to be unwarranted to remove a criterion based upon such judgments.

Similarly, the suggestion that some of the poorer scoring criteria should be redefined is also criticised. In fact, Tully even suggests giving raters broader latitude in their interpretations of the CBCA criteria. He compares SVA to a clinical examination in this regard; protocol not instrument. Under this conceptualisation, Tully states that, "[r]eliability arises from the valid principles which guide CBCA as part of SVA, and not from its procedural standardisation per se" (p.186).

Horowitz (1998) replies to this critique in the same issue. In agreement with Tully, he concedes that the quality of the transcripts does affect reliability and that it was premature to suggest that some of the criteria be dropped. However, he suggests that transcript quality is only one of a number of untested facets of generalisation that were not scrutinised in the original study. The particular sample was selected in order to maintain homogeneity and all that was needed was for the study to be replicated. Nevertheless, Horowitz sees no reason to reappraise the psychometric goal of SVA, and warns that, "we should not sink to a blind and romantic defence of human, clinical decisions" (p.190). He concludes by saying:

"That CBCA has not yet met the psychometric standards should spur us to more research. It should not leave us to touting the method of authority for establishing the credibility of a statement" (p.190)

The argument over whether or not CBCA should be psychometric instrument or part of a protocol reveals an uneasiness concerning the theoretical underpinnings of the technique. Reverting SVA towards the status of a protocol in many ways narrows its horizons and makes the technique easier to legitimise. Nonetheless, even if it does not attain psychometric standards at first, the reliability of CBCA can only be helped by pushing its development towards a more standardised approach. It seems probable that the catalogue of criteria

provided by CBCA will need future modification and that this may prove difficult without a strong theoretical grounding on which criteria may be derived.

### **3.1.5.6 Realism and Construct Validity**

The most pertinent critique of experimental SVA studies is that they lack mundane realism and that the 'lie' scenario does not appropriately match the scope of the phenomena of interest. A key concern is whether participants are lying when they are told to do so. In certain studies, participants are told to relate personal stories that are traumatic, involve loss of control, and involve negative emotions (Landry and Brigham, 1992; Ruby and Brigham, 1998). Similarly, Sporer (1997) asked for an event that was freely invented, not taken from a book or a friend, and that was not a mere modification of another event experienced by the participant. All of these authors acknowledge that it is difficult to verify whether or not participants were actually lying or indeed, whether they were each lying to the same degree. Ruby and Brigham (1998) concede that this is an inevitable drawback of this type of research. There is probably little reason for a significant number of participants to sabotage experiments in this manner, and so this risk to realism could be deemed acceptable.

Some investigators have taken a different approach whereby the truth-status of the story is known already (Köhnken et al., 1995; Akehurst et al., 2001; Zaparniak et al., 1995; Vrij et al., 2000). Participants are usually asked to either watch a videotape or listen to a verbal description that serves as the subject matter for a subsequent interview. Participants in the lying condition have to pretend that they were actually present at the event. This technique has the advantage of ensuring that all of the stories are similar in type. It might be argued however that such lying behaviour is too subtle, embodying the 'story modification' approach guarded against by Sporer (1997). However, from the perspective of external validity this may not be a concern because real suspects may quite likely use televised material as a basis for their lies (Akehurst et al., 2001).

Another concern is that this approach lacks the personal relevance that is gained from the telling of traumatic personal experiences and so the participant will not be highly motivated to lie. In one study with nursing students (Vrij et al., 2000), motivation was enhanced by telling the students that being able to lie effectively was an essential part of their future careers. Personal relevance can also be increased by utilising a 'mock crime' scenario, typical in other types of lie-detection study. In the study by Porter and Yuille (1996), participants engaged in a mock crime under the experimental pretense of testing an in-house security contractor. This then served as the interview topic and the participants were informed that a reward of extra money would be given to them if they actually succeeded in 'fooling' the interviewer.

The truth status of the transcripts is foremost amongst the construct validity concerns of field studies. Court rulings are fallible and are often subject to an innocence (truth) bias (Ekman, 2001) and so they are not a sufficiently objective measure against which to configure CBCA. Sometimes only cases where a full confession has been made by the perpetrator are

considered (Anson et al., 1993; Parker and Brown, 2000) but even confessions can be unreliable. As mentioned earlier, this was addressed by Lamb and his colleagues (1997) who used ICFSs to provide a variety of information from which to assess the statements of allegedly abused children.

### **3.1.5.7 External Validity**

SVA and CBCA were developed with children in mind. One of the issues concerning researchers is the extent to which CBCA is affected by the culture, gender and age of the witness. To what extent can the SVA technique be applied outside of the domain of child sexual abuse, and possibly outside the forensic domain altogether? SVA is grounded upon the Undeutsch hypothesis which itself is based around qualitative and quantitative distinctions between true and false statements. To what extent do the qualitative and quantitative distinctions across other individual differences overlap with these veracity distinctions?

So far, there has been little evidence to support the hypothesis that gender confounds total CBCA scores (Lamb et al., 1997; Landry and Brigham, 1992; Sporer, 1997). Nonetheless, Sporer (1997) did find some univariate effects with individual criteria. For example female targets tended to make more references to psychological processes.

There has been some evidence that shows a correlation between total CBCA scores and the age of preadolescents (Lamb et al., 1997; Anson et al., 1993). This effect has not been demonstrated for children above a certain age and is almost certainly correlated with a child's cognitive development. Anson et al. (1993) found a significant positive correlation between age and several of the criteria (c1, c4, c5, c6, c18, and c19). Parker and Brown (2000) found no significant influence of age on total CBCA scores but did mention that the patterns of scores were different. For example, they cite c10 ('Accurately Reported Details Misunderstood') as less likely to occur with adults who are more likely to understand details about the offence. This may be true in the particular case of sexual offences but also in other domains as well. For example, there are likely to be various bureaucratic and technical details that can be reported by some adults without a full understanding of the process that is occurring.

There have been few studies that have examined the cross-cultural impact of CBCA. In fact, there has been little research performed outside of Germany, the UK, and the USA altogether. Lamb et al. (1997) performed their research with Israelis using the Hebrew language and managed to find a correlation with CBCA and truth status, but this was not a cross-cultural examination. Ruby and Brigham (1998) tested the relationship between CBCA scores and target ethnicity of Black (Afrocentric condition) and White (Eurocentric condition) participants. They observed that in the truthful transcripts, c3, c6, c12, c14, and c17 were rated more often for Black targets than they were for White targets. Additionally, when the results were combined for *both* true and false accounts, c1 and c4 occurred more often in the statements of White targets and criteria c2, c5, c6, c8, c9, c11, c12, c14, c15, and c16 occurred more often in the statements of Black targets. It is too early to assess the impact of these results, but from an

explorative viewpoint, they do suggest that the extent of CBCA's cultural relevance needs to be investigated further.

### **3.1.5.8 Criteria Weighting and Configuration**

There exists a variation in the discriminatory power of individual criteria across studies. For example, c2 and c3 are often seen to be highly effective discriminators whereas c10 is not <sup>3</sup>. This has led some researchers to suggest that perhaps some criteria are more valid than others, or are at least more relevant in certain circumstances (Steller and Köhnken, 1989). Weighting and configuration systems (such as the decision rules proposed by Zaparniak et al., 1995) have been proposed to boost the relevance of certain criteria. The relationship between the relevance of certain criteria over others and the recounting of particular types of episode remains a subject for future research.

### **3.1.5.9 New Criteria**

It seems almost inevitable that an empirically derived technique such as CBCA should become augmented with additional criteria once the theoretical foundations have been explored more fully. Köhnken's study examining the effect of the cognitive interview on CBCA investigated the following six additional criteria (Köhnken et al., 1995). These criteria were derived from an earlier unpublished study.

1. *Reporting Style (long winded vs. issue-related)*
2. *Display of Insecurities*
3. *Providing Reasons for Lack of Memory*
4. *Clichés*
5. *Repetitions*
6. *Comments and Interpretations Regarding the Event*

All of the criteria were found to significantly discriminate between the truthful and fabricated accounts with the exception of criteria six. 'Reporting Style', 'Insecurities', and 'Reasons for Lack of Memory' were especially significant ( $p < 0.01$ ). These last two are similar to the CBCA criteria of c15 and c16 ('Admission of lack of memory', 'Displaying doubt about testimony') and probably could be incorporated into the existing definitions, whereas the 'Reporting Style' criterion is novel.

Given that CBCA is far from reaching extremely high levels of accuracy, there always remains the scope for expanding the technique. This could involve producing new verbal criteria or complementing the analysis with other approaches. There may be a natural ceiling for just how accurate CBCA can go by itself. If this is the case, then the evolution of this route to lie-

---

<sup>3</sup> C.f. meta-analyses of the effectiveness of individual criteria in (Vrij, 2000; Ruby and Brigham, 1997)

detection may ultimately hinge upon a multi-channel approach that also looks at nonverbal and physiological clues.

## 3.2 Reality Monitoring

The Reality Monitoring (RM) technique provides an alternative to CBCA rather than the entire SVA process. Like CBCA, it is a criteria-driven approach except that it has the advantage of being theoretically derived and relatively easy to teach to raters. The RM approach is still very much in its infancy but this section will discuss what advances have been made in developing the technique for truth-validation.

### 3.2.1 Background

If memory is conceptualised as both a constructive and reconstructive cognitive system, then it is reasonable to hypothesise that such reconstruction may lead to potentially distorted memory traces (Mitchell and Johnson, 2000). The Source Monitoring framework (Johnson et al., 1993) and the more specific Reality Monitoring model (Johnson and Raye, 1981) were developed in order to specify how this system might still remain functional, and how a person is able to attribute the origin or the source of their memories.

It has been said that memories that are false and those that are true are often products of the same underlying cognitive mechanisms (Johnson and Raye, 1998). Johnson and Raye (1998) envisioned a cognitive system that focused upon the contents of the memories themselves. They suggested that memories might be attributed to a particular source based upon two evaluation processes. Primarily, a low-level heuristic mechanism would base attribution on the degree to which a particular memory engram fitted a fuzzy set of characteristics identifying a particular source. A higher-level, slower process may then employ biases that attribute the engram to either the same or a different source.

Johnson and Raye's (1981) RM model posits two sources of memory engram, those derived from external or 'real' perceptual information, and those derived from self generated information such as re-representation, co-temporal thoughts, or fantasy. Attribution processes would then discriminate between these sources based upon four characteristics of the engram content.

1. *Spatial and Temporal Detail*
2. *Sensory Attributes*
3. *Semantic Detail*
4. *Detail Regarding Cognitive Operations*

More specifically, externally generated memories would exhibit more of these characteristics than internally generated memories with the exception of the last criteria. As perception is for the most part an automatic process, it was hypothesised that less information would be available that contained references to cognitive operations in comparison to an internally

generated memory, where information concerning these autonomic processes would need to be made explicit.

The role of RM in content analysis was to provide the missing theoretical basis for why the internally generated stories of a deceiver were qualitatively different from those externally generated by a truth-teller.

### **3.2.2 The Criteria**

One set of criteria that was based upon the RM model was defined by Sporer (1997). Sporer described how the development of a lie-detection technique based upon RM would require a research programme that was very different to that of existing RM research. Firstly, the aim was not to address how people made the reality distinction internally, as liars already know the reality status of their accounts, but whether outside judges could do the same using similar criteria. Secondly, the technique does not concern itself with memories, but the operationalisation of those memories through verbal descriptions (Sporer, 1997; Vrij, 2000).

Sporer defines eight categories that are based upon the four criteria presented by Johnson and Raye (1981), and on previous studies in which Sporer was involved concerning the 'Judgments of Memory Characteristics Questionnaire' (Sporer and Kuepper, 1995; Sporer, et al., 1995).

1. *Clarity*
2. *Sensory Experiences*
3. *Spatial Information*
4. *Time Information*
5. *Emotions and Feelings*
6. *Re-constructability of the Story*
7. *Realism*
8. *Cognitive Operations*

The first seven of these criteria concern truth-validation whilst the eighth concerns lie-detection.

### **3.2.3 Empirical Evidence**

Sporer (1997) undertook a study where forty undergraduate students were recorded telling both one truthful story and one complete fabrication. This yielded eighty transcripts that were coded by two reviewers trained in the RM scoring method. The overall accuracy of the technique was scored at 71.3% with a slight higher accuracy rate of 75% for the true stories. A further factor analysis was also performed with the scores yielded from CBCA ratings on the same transcripts. This analysis yielded five factors of (1) Logical Consistency and Realism, (2) Quantity of Details and Contextual Embedding, (3) Feelings and Thoughts, (4) Clarity, and (5) Verbal and Nonverbal Interactions. Sporer hoped that the identification of such common factors between CBCA and RM would lead to the emergence of an "integrative cognitive-social theory of the detection of deception" (p.392).

Vrij and his colleagues (2000) have also investigated the effectiveness of the RM criteria, again comparing the results with CBCA. In their study, seventy-three nursing students were told to either lie or to tell the truth when asked questions about a videotape that they had viewed. Again, two raters were trained on the RM criteria. It was found, in agreement with Sporer that it was easier to instruct the interviewees in the RM criteria as opposed to the CBCA criteria. This is a clear advantage of RM over CBCA, providing a higher level of inter-rater reliability. A modified set of criteria was used that incorporated the categories (1) Visual Perceptual Information, (2) Audio Perceptual Information, (3) Spatial Information, (4) Temporal Information, and (5) Affective Information. The category of 'Cognitive Operations' was dropped as it had yielded poor discriminatory power in previous research (c.f. Alonso-Quecuty 1992,1996; Höfer et al., 1996) and because it was a lie-detection and not a truth-validation criterion. The study yielded accuracy rates of 67.12% overall with 70.6% of truthful statements and 64.1% of fabricated statements classified correctly (CBCA scored 72.6% overall, 64.7% for true and 79.5% for fabricated).

### 3.3 Other Verbal Approaches

To date, most of the contemporary research into the verbal discriminates of deception has focused upon SVA and Reality Monitoring. For completeness there are two other techniques that are also relevant. The first of these is a commercial technique called SCAN and the other examines the lexical diversity of a statement.

#### 3.3.1 SCAN

Scientific Content Analysis (SCAN) is a technique that is developed commercially by 'The Laboratory for Scientific Interrogation, Inc.' (Sapir, 1987). This technique follows a similar format to RM and SVA, a data elicitation followed by a content analysis. However, the elicitation method involves getting the witness to write their statement first before it is subjected to analysis. LSI has also developed a special questionnaire to give to the witness (VIEW) that can be used in place of the written statement. The principles underlying SCAN are apparently based on a witness' written 'impression management' as apposed to the more impulsive, unrehearsed statement preferred by SVA. The developers of SCAN claim scientific validity scores of up to 95% "supported by a validation survey conducted in a U.S. governmental agency" [<http://www.lsiscan.com/id17.htm>]. Nonetheless, there has been little published material investigating these claims.

One study that attempted this sort of validation was conducted by Porter and Yuille (1996). The study (a typical SVA-type scenario) compared CBCA and RM criteria along with type-token ratios (see the next section) and three computer-generated criteria taken from SCAN. These criteria were operationalised as:

- *Frequency of Unnecessary Connectors – For example, connectors that aim to 'pass-over' incriminating information such as 'the next thing I remember'.*



- *Frequency of Pronoun Usage Deviation* – A tally of how many times the witness deviated from the first person and the past tense
- *Ratio of Component Elements* – The ratio of introduction length to conclusion length, calculated in 'Number of Words'

Porter and Yuille (1996) found no discriminatory power of these criteria in determining truthful statements and warn that, "training programmes such as Sapir's may not be disseminating wholly accurate information to investigators" (p.453). Interestingly, out of all of the criteria analysed they only managed to find support for three (all CBCA). They mention that more work of this kind is needed before SCAN could be dismissed out of hand. This is especially the case given the small sample and potential construct validity pitfalls of how the SCAN variables were operationalised. Shuy (1998) defends SCAN further by highlighting the point that "in all fairness to SCAN, no claim is made that subjects whose statements are judged to have produced deceptive language are actually lying. [...] SCAN uses such language as a guide to the interviewer about what to probe, when to challenge, and what to disbelieve in the follow-up interrogation" (p.105). This is not much of a defense, since it promotes an interrogative style based upon invalidated clues. This could lead to the trap of self-feeding misperceptions and Othello errors. Ekman (2001) begins his well-respected book *Telling Lies* stating his concern with so-called "'experts' who go unchallenged by public scrutiny and the carping critics of the scientific community". In sum, there has been insufficient scientific analysis (probably due to commercial reasons) to ascribe SCAN's 'intuitively appealing' criteria (Porter and Yuille, 1996) any foundation of empirical validity.

### **3.3.2 Lexical Diversity**

Porter and Yuille (1996) also investigated the relationship between an eyewitness' lexical diversity and the veracity of their statements. Lexical diversity is commonly investigated using the type-token ratio (Porter and Yuille, 1995). This ratio compares the number of unique words against the total number of words in a statement. It has been suggested that motivation to appear credible might reduce the range of vocabulary used by an eyewitness, making their speech more stereotypic (Hollien, 1990). Alternatively, it has also been demonstrated in a case study by Carpenter (1990) that the increased vigilance that a deceiver has over their responses might actually increase lexical diversity as statements are chosen more carefully.

In the same study in which SCAN was investigated, Porter and Yuille (1996) compared the type-token ratios across all four of their experimental conditions. There was no more than a one percent difference between the average type-token ratios measured. Porter and Yuille guard against dismissing lexical diversity as a reliable indicator on the basis of this one study, and it is interesting to note that the study yielded very little discriminatory power for any of the other criteria used. In addition, a more recent study by Zhou et al. (2003) reports deceivers having a lower lexical diversity when communicating over email. Nonetheless, it seems likely that trying

to diagnose deceit based on type-token ratios alone is overly reductionist and is probably not a very promising route to lie-detection.

### 3.4 The Polygraph

The polygraph has been the subject of a great deal of controversy but nonetheless has still persisted as the most widely-used lie-detection technique. The polygraph is used by the courts in some US states, by personnel departments for the purposes of employee screening, and even by the hosts of talk-shows in order to 'detect' cases of infidelity for the viewing public. Companies exist that provide polygraph examinations, and polygraph examiners in the US have their own organisational body (the American Polygraph Association). Despite this popularity, the polygraph has received comparatively little scientific scrutiny. In this section, literature pertaining to two of the most popular polygraph tests, the Control Questions Test (CQT) and the Guilty Knowledge Test (GKT) will be discussed <sup>4</sup>. The principal proponents of each technique, David Raskin (CQT) and David Lykken (GKT) have taken adversary stances within the literature, such that the relative merits of the two methods have been debated at length. This debate will be presented here.

#### 3.4.1 The Control Questions Test

A typical CQT as used by law enforcement officers comprises of three stages <sup>5</sup>. Initially, the examiner must read up on the case-file for the crime in question and then a pretest interview is held with the suspect. The purpose of the interview is for the suspect to relate their side of the story to the examiner, and for the examiner to develop a series of 'yes or no' questions that will be used during the test. These questions are then discussed with the suspect so that they can understand them and are able to give a definite yes or no answer for each one. The examiner is then ready to administer the test.

The CQT comprises of three different types of question (relevant, irrelevant, and control). Relevant questions are directly related to the incident, such as *"Were you at the crime scene on the night in question?"*. Control questions are questions that are similar in nature to the relevants but that are unrelated to the crime. These questions are designed in a way that causes the examinee to become anxious about the morality of their past history, encouraging them to lie. An example of a control question would be something like *"Have you ever stolen something from someone?"* which, in its broadest sense, might be true of anyone. If the suspect responds in the affirmative, the examiner may then narrow the scope of the question, explicitly increasing the moral weighting until a suitable control question (i.e. a 'no' response) is found. For example, the previous question may be narrowed to the inquiry *"But have you ever stolen something from someone who really trusted you?"* Irrelevant questions are innocuous and do

---

<sup>4</sup> Literature concerning the older Comparison Question Technique and the Relevant-Irrelevant Test has been omitted from this review.

<sup>5</sup> The description given by (Ben-Shakhar, 2002) forms the basis for this account.

not relate to the incident in question. Usually they are used to absorb an initial orienting behaviour to the first question, and to provide a gap between the control-question / relevant-question pairs. An examination of the polygraph charts would typically compare control-responses against relevant-responses. The underlying theory behind the CQT is that responses to control questions will be more pronounced than that of relevant questions for truthful examinees and vice versa for deceptive ones.

There have been a few, limited attempts to scientifically examine the validity and reliability of the CQT (see Ekman, 2001). Some of this effort has been driven by John Reid, the original creator of the CQT (Reid, 1947; Reid and Inbau, 1977). In a 1971 article, Horvath and Reid attempted to demonstrate the effectiveness of the CQT using data collected from actual investigations conducted by the second author. Ten examiners (employees of John E. Reid & Associates) examined forty polygraph tests without being given access to accompanying case files. The seven 'experienced' examiners attained accuracy levels of 91.4% and the three less experienced examiners achieved 79.1%. These encouraging results are somewhat tempered by the selectivity of the sample. Other studies have only achieved similar accuracy levels when 'inconclusive' results are omitted. Horvath, in a later study (1977) used law enforcement officials as 'blind' raters in the analysis of a stratified, random sample of 112 polygraph records. An overall accuracy level of 63.1% was achieved, with higher levels of accuracy for deceptive examinees than for truthful ones (Horvath and Reid found the opposite effect; 1971). Confessions were used as an indication of ground truth and Horvath acknowledges the possibility of a sample bias towards 'confessors' who may differ affectively from 'non-confessors'. Barland and Raskin (1975) attempted to test CQT against a valid measure of ground truth by using a mock-crime scenario that involved participants stealing \$10 from the drawer of a desk. Motivation was operationalised by promising the participants (undergraduate students) that they could keep the money if they passed the polygraph test. Only 53% accuracy levels were achieved by six examiners but this figure rose to 81% when the 'inconclusive' polygraph reports were removed from the sample. In another study by Raskin, accuracy levels of 88% were achieved with this rising to 96% when inconclusives were dropped from the analysis (see Podlesny and Raskin, 1977). David Raskin has become the leading advocate of the CQT technique.

The validity of the CQT has come under considerable fire from David Lykken, an accomplished psychophysicologist (Lykken, 1974, 1979). Lykken (1979) has stated that, "the theory of the lie test is so naïve and implausible that one should demand especially strong empirical evidence before accepting claims of extremely high validity" (p.47). It would seem that studies such as the ones mentioned so far, fail to live up to these expectations. Lykken suggests that even if 'objective and consistent rules' for interpreting complex polygraph amplitude measures *were* available (the implication being that practically they are not), there is still a considerable difficulty in determining the nature of the responses to control questions. There is no way to validate whether a person is actually lying to these questions, or more specifically, believes that they are lying. Lykken suggests that the examinee might often interpret these questions in such a way so

that their responses are perceived as being truthful. For example, the question *"Have you ever stolen something from someone?"* may be interpreted as only referring to a criminal offense. Lykken challenges the status of control questions in providing a control "in the usual scientific sense of the term" (p.49). He also suggests that there is an inherent bias towards elevated autonomic responses to relevant questions, regardless of guilt status, because of the fact that they are indeed relevant "to the source of [the examinees] immediate jeopardy" (p.49).

Lykken also attacks the empirical support for the CQT. He dismisses mock-crime studies on the grounds of poor psychological realism, and any study whereby examiner's judgments were not derived entirely from polygraph reports. More rigorous studies such as Horvath's (1977), with lower reported accuracy and a relatively high rate of false positives (39%) tended to show a more accurate picture of the effectiveness of the polygraph. Lykken goes on to speculate on the effectiveness of countermeasures. Lykken (1979) suggests that a CQT might be beaten by an examinee flexing their toes, tensing their diaphragm, or biting their tongue at particular moments during the test. He lays down the challenge that he could train 'guilty' people to beat the CQT:

"I claim that I could train guilty suspects to successfully 'beat' the control question lie test [...] let me train guilty suspects in Barland and Raskin's next field study, and I predict their false negative rate may approach what their false positive rate is right now" (p.53).

In a direct response to Lykken's critique, Raskin and Podlesny (1979) accuse Lykken of failing to understand "the simple, basic theory" (p.54) concerning the evaluation of relevant and control responses. They claim that the purpose of the control items is not to provide a scientific control, but a stimulus that will identify innocent examinees by eliciting stronger autonomic responses. Further, they object to Lykken's dismissal of laboratory studies, emphasising the attempts made to ensure realism and proclaiming that such controlled approaches are central to scientific method. The results from the Horvath (1977) study are qualified on the grounds that the examiners were relatively inexperienced. Further, they present an analysis of five additional studies, performed with blind raters that demonstrate high levels of accuracy. With respect to countermeasures, Raskin and Podlesny refer to an earlier study, conducted by Lykken himself (1960) that showed the poor effectiveness of countermeasures with regard to the Guilty-Knowledge Test.

Subsequent studies have attempted to address some of the criticisms leveled by Lykken concerning scientific CQT research, with results that are less encouraging than those reported within Raskin and Podlesny's (1977) article. One field study that managed to attain ground truth within a realistic scenario was conducted on Israeli police officers by Ginton and colleagues (1982). Officers were given a test as part of a training course and then were allowed to correct their own answers. A chemical sheet, underlying the answer paper was used to record whether the answers were subsequently altered. Seven officers cheated out of a total of twenty-one. Some of these later confessed when suspicions were raised, but two of these (along with

thirteen others) agreed to take part in a CQT. Initial evaluations correctly identified one of the two guilty officers, and seven of the thirteen innocent officers.

Modern research and analysis has led to increasing support for Lykken's arguments against the CQT. A full review of this literature is beyond the rather selective scope of this chapter. A modern summary of the flaws of the CQT, undermining its scientific credibility has been presented by Ben-Shakhar (2002).

1. *There is a lack of equivalence between Relevant Questions and Controls and this systematic bias is liable to lead to increased false positives.*
2. *The CQT is not a standardised test.*
3. *The CQT does not rely on objective criteria for interpreting the polygraph charts.*
4. *The CQT is prone to contamination by subjective impressions formed by the examiner in reading the case literature*
5. *It has turned out that the CQT is actually quite vulnerable to countermeasures.*
6. *There is an absence of sound scientific research that lends support to either the CQTs validity or reliability.*

The CQT it would seem, may have utility in extracting confessions from suspects, however its claims to validity are currently very suspect (Iacono and Lykken, 1999).

### **3.4.2 The Guilty-Knowledge Test**

The Guilty Knowledge Test (Lykken 1959, 1960) is designed to establish whether a suspect has knowledge of a crime that they would not recognise if they were innocent. This sort of knowledge would be restricted to information that has not yet been reported in the media or otherwise become widely known. A hypothetical scenario can be assumed where the police have found a murder weapon at a crime scene (a particular type of pistol). The GKT test may involve showing the suspect a series of pictures of different types of handgun, and each time the suspect did not recognise the particular gun they were to press a button indicating 'no'. Similarly, the test may be conducted verbally, whereby the suspect is asked a series of questions such as "Was the murder weapon a Colt .44 calibre?" to which the suspect would reply, "I don't know". Within the series of pictures or verbal descriptions, the actual murder weapon would be inserted. The central hypothesis underlying the GKT is that a guilty suspect would autonomically recognise the murder weapon, and this brief recognition would then be distinguishable from the responses to other stimuli. The test would be repeated over a selection of different questions enhancing the probability that the recorded response was related to the attainment of 'guilty knowledge'.

Lykken proposed that the GKT is the only scientifically grounded (and credible) polygraph test. The advantage of the GKT, is that the phenomenon under detection is a well-defined response that is well known to scientists. The 'recognition' response evoked during the test is known as the orienting reflex (Ben-Shakhar and Elaad, 2002). This is an automatic reflex that orients a perceiver to a novel change in stimuli (in this case the appearance of the murder weapon). The

nature of the orienting reflex (OR) requires that the stimulus types be chosen with great care so that it is clear that the OR has not been evoked because of some other feature of the stimuli. For example, if the murder weapon (a gun) was placed in amongst series of different knives, then an alternative hypothesis would suggest that the OR was evoked as a result of the novel stimulus of a 'gun' appearing amongst 'knives'.

The GKT has received relatively strong support from empirical studies. In a recent review of fifteen mock-crime studies, accuracy levels were estimated overall at an 80.6% correct classification rate for innocent suspects and 95.9% rate for guilty suspects (Elaad, 1998). In addition, only four of these fifteen studies contained false-positive errors (innocent suspects diagnosed as guilty). Similarly, a review of ten laboratory studies by Ben-Shakhar and Furedy (1990) reported accuracy rates of 94.2% for innocent suspects and 83.9% for guilty suspects respectively. Therefore, it would seem that the GKT is a promising form of polygraph testing.

Despite this, the GKT is limited to the scenarios in which it can be used. One of the key problems in the field is the construction of appropriate guilty-knowledge questions (Ben-Shakhar and Elaad, 2002). A successful GKT typically requires four or more guilty knowledge questions to be used (Podlesny, 1993) with greater accuracy being attributed to tests with a larger number of questions (Ben-Shakhar and Furedy, 1990). Even so, an investigator must be careful to design the test so that a large number of repetitions do not cause a process of habituation that dampens the OR (see Siddle et al., 1983).

As with the CQT, traditional polygraph dependent measures may be susceptible to manipulation through countermeasures. However, because the GKT involves the detection of the OR, the electroencephalograph (EEG) can also be used. ORs are exhibited as P300 Event-Related Potentials (ERP) after the activation of a novel stimulus (Siddle and Packer, 1991). These ERPs are less susceptible to countermeasures than other polygraph measures. However, the P300 trace is also produced by other reflexes as well as the OR, impressing the need for careful experimental control.

### **3.5 Summary**

In this chapter, the techniques of Statement Validity Assessment, Reality Monitoring, SCAN, and Lexical Diversity were presented as the verbal routes to lie-detection. SVA whilst being the most successful of these is still far from being highly deterministic, and it would seem that the technique needs further development before it can be used with confidence within a forensic setting. The main hindrance to the development of SVA lies in its empirical grounding. Techniques such as Reality Monitoring have provided alternative, more theoretically derived coding systems that might be plugged into SVA in order to enhance its validity. The commercial technique SCAN, as with many other commercial lie-detection products appears to make claims unjustified within the academic community. The lexical diversity approach has yet to be shown to be highly successful and is in danger of straying to close to an inappropriate reductionism. In section 3.4, the two most popular polygraph tests were reviewed; the CQT and the GKT. Of

these, the GKT has received the most credible scientific support, despite not being as widely applicable as the CQT.

It may be possible that existing verbal techniques such as CBCA can be applied to statements attained using CMC. Physiological responses such as the orienting reflex may similarly be captured through the low-level analysis of keyboard and mouse behaviour. Nonetheless, communication through CMC differs both quantitatively and qualitatively from face-to-face communication. Any lie-detection strategy that uses these technologies must factor such differences into the technique. The research that has examined these differences has been provided in the next chapter and comprises the final part of this literature review.

## Chapter 4

# Computer-Mediated Communication

---

Over the last decade, the breadth of CMC research has grown steadily and it is now difficult to consider the discipline as a single field of inquiry. CMC technologies have been studied within a variety of contexts including psychotherapy, distance-learning, computer-supported cooperative work, and human-computer interaction. Online communities have been studied at length along various sociological dimensions, and investigators of qualitative research methodologies have used CMC to provide them with new tools and techniques.

A selection of CMC research is presented within this chapter. This review will focus upon on those areas that are directly applicable to the study of interpersonal deceit. This treatment begins with a brief section presenting the history of CMC technology and research, in order to provide a necessary background to the domain and the theories presented throughout the rest of this chapter. The principal theoretical developments concerning mediated communication are then discussed, including Social Presence Theory, Hyperpersonality Theory, and emergent approaches such as the Gibsonian Ecological Perspective. The structure of one particular type of CMC (quasi-synchronous chat) is then examined at a lower level of abstraction, revealing the micro-level characteristics and affordances of this particular discursive form. Identity deception is one of the most prevalent forms of online deceit and therefore the construction of the online 'self' will be considered towards the end of the chapter. Finally, the handful of studies that have explicitly examined deception or lie-detection within CMC will be reviewed.

### 4.1 The Origins of CMC Technology and Research

On the 24<sup>th</sup> May 1844, a simple, encoded message is being transmitted sixty kilometers along a cable that lies between the US cities of Washington and Baltimore<sup>1</sup>. The message, "What hath God wrought!" is traveling along the world's first telegraph line and is signifying the birth of telecommunications. The telegram, a form of text-based communication, was not simply an electronic letter, but a medium that was associated with its own unique style of discourse. This style was born of both technological *and* sociological concerns. For example, senders would often abbreviate or omit words within telegrams in order to keep the messages shorter and therefore cheaper. Later, in 1876, Alexander Bell invented the telephone. This medium also brought with it a change of communicative style. Visual anonymity demanded an alternative form of conversational regulation that did not depend upon nonverbal cues. As it will be

---

<sup>1</sup> Unless otherwise referenced, the historical information pertaining to the development of CMC technology has been sourced from (Moschovitz et al., 1999).



demonstrated throughout this chapter, text-based CMC shares these socio-technological characteristics with the telegram and the telephone.

The year 1958 sees the development of the first modem by Bell labs, and by 1969, the ARPAnet computer network (the precursor to the Internet) is connecting four research sites across the US. The 1970's see the birth of CMC. In 1971, Ray Tomlinson has produced the first email program and has conventionalised the use of the '@' symbol in email addresses. In 1973, David Wooley produces PLATO Notes, an early form of 'Bulletin-board Service' (BBS). This system allowed users to post emails to a central 'board' where they would be arranged in sequence according to a particular topic (thread). The first public BBS (the Computerised Bulletin Board Service) goes online in 1978, and provides the public with one of the earliest forms of online community. For the first time, public users can develop interpersonal relationships and discuss topics with other people that they have met solely through the CBBS system.

Even as early as the late 1970s, researchers are already beginning to examine the possibilities afforded by email, BBS, and other 'computer-conferencing' systems. This early research was undertaken in a technologically deterministic style that focused primarily on task-oriented communication. Most notable is the research on the EIES system performed by Star-Roxanne Hiltz and Murray Turoff (an early CMC pioneer himself) summarised in their seminal book 'The Network Nation' (1978). Another early, but less well-known book 'Electronic Meetings: Technical Alternatives and Social Choices' by Johansen, Vallee, and Spangler (1979) also looked at the benefits (and weaknesses) of computer-conferencing.

The late 1970s also saw the beginnings of Usenet, a BBS-style system whereby users could subscribe to different 'channels' that were hierarchically arranged by topic. People could then post to the channels and view all of the postings according to thread. Online communities within Usenet soon started to adopt various conventions from which many of the modern day Internet conventions are derived. For example 'smileys', the practice of typographically representing faces to depict emotion, was allegedly started by Kevin Mackenzie in 1979 from an idea gleaned from 'The Readers Digest' (Moschovitis et al., 1999). Nowadays, the use of smileys and other 'emoticons' are commonplace. The year 1979 also saw the creation of the first 'Multi-User Domain/Dungeon' (MUD) by Roy Trubshaw and Richard Barthe, students at Essex University. MUDs represent the merger of text-based adventure gaming with real-time text-based chat, to produce a multi-player virtual environment in which people can create and explore a world depicted wholly through text. MUDs and their successors MOOs (MUD object-oriented) became increasingly popular throughout the 1980's.

The 1980's saw the continued development of CMC technologies and communities based upon MUDs, BBSs, Usenet and Email systems. Researchers investigating the social-psychological aspects of these systems were beginning to conduct research under what was later to be called the 'filtered-cues' paradigm. Most famous of these researchers were the group based at Carnegie-Mellon University, who produced the seminal article, 'Social Psychological Aspects of

Computer-Mediated Communication' (Kiesler et al., 1984) and other related research on task-based group interaction within the medium (much of which is reviewed in the book 'Connections' published by the same team; Sproull and Kiesler, 1991). Other researchers, adopting a qualitative style were starting to examine some of the newly flourishing online communities such as CompuServe, the Well, and Lucasfilm's Habitat. One popular account is provided by the author Howard Rheingold (1993) who describes his experiences in 'The Well' in the book 'The Virtual Community'. These accounts describe emergent phenomena that have seeded within these communities, including flaming (the act of posting abusive, aggressive emails), pseudonymity, textual conventions, and online identity exploration. So-called 'cyberpunk' novels are appearing that explore the potential of online communities such as William Gibson's (1984) 'Neuromancer' in which the term 'cyberspace' first appears. In 1983, the Internet emerges out of ARPAnet.

The 1990s have seen the popularisation of real-time text-based chat programs. Systems have emerged such as Internet Relay Chat (IRC), ICQ, and web-based chat programs. Modern-day virtual communities such as Active-Worlds have taken MUD-like universes and provided them with graphics, creating fully navigational three-dimensional environments.

Research in the 1990s began to focus away from the filtered-cues paradigm and towards a relational approach. Postmodern research, inspired by the online community phenomenon has proliferated. The last few years have seen an increased focus on the methodology and the ethics of studying online communities. CMC research is now a very broad field covering a vast number of disciplines (however this breadth of research often comes at the sacrifice of depth). Similarly, the number of different types of CMC system has grown making it difficult to meaningfully refer to CMC research under a single banner title. As was stated in the introduction, the research reviewed throughout this chapter will not attempt to survey the entire breadth of this banner, but will instead focus on those areas most pertinent to the study of lies and deceit. This review will begin, back in the days of ARPAnet, a few years after Tomlinson delivered that first email message in 1971.

## **4.2 Theoretical Approaches to CMC Research**

### **4.2.1 The Social Presence Approach**

The earliest theoretical approach applied to CMC was described by Short, Williams and Christie (1976) in their book 'The Social-Psychology of Telecommunications'. Their Social Presence approach has become something of a staple antithesis in contemporary CMC literature reviews. Although CMC is mentioned in passing (this was long before the technology had become widespread) the authors focused primarily on other telecommunications media such as the telephone. They were rather more tentative in their theorising than might initially be assumed from the CMC literature-base. This was the original theory under which CMC could be classified as an inherently impersonal, task-oriented, 'cold' medium.

It was proposed by Morley and Stephenson (1969) that the two functions of communication, the interpersonal and the 'interparty', exist in a balance whereby the medium might facilitate one function at the expense of another. According to Short, Williams, and Christie (1976), a visually deprived medium such as the telephone might then facilitate the interparty function at the expense of the interpersonal. They define 'social presence' as a perceived quality of a telecommunications medium that reflects "the degree of salience of the other person in the interaction, and the consequent salience of the interpersonal relationship" (p.65). Social presence affords a certain type of interaction to a particular user or type of user rather than being a solely objective quality of the medium itself. The individual is to some extent aware of the social presence of a particular medium and will preferentially select it based upon the degree to which they require interpersonal or task-oriented interaction.

By this reasoning, a medium such as text-based CMC would typically score very low in terms of social presence and would be more desirable in circumstances where task-oriented interaction was key. In a world before virtual communities, where CMC was used predominantly in institutional, task-oriented settings, this hypothesis would have seemed a lot more reasonable. The concept of social presence was never really elaborated much beyond the 1976 formulation and has been a source of some confusion amongst researchers (Steinfeld, 1986; Hiemstra, 1982; Svenning and Ruchinskas, 1984; Walther, 1992).

### **4.2.2 The Filtered-Cues Paradigm**

The Filtered-Cues paradigm is a general perspective that, like Social Presence theory, was derived from early ideas about the social 'bandwidth' of media. It was hypothesised that a medium that allowed less channels of communication such as visual, verbal, tactile, or auditory channels would hinder the passage of relevant interpersonal information (Wiener and Mehrabian, 1968; Weston and Kristen, 1973)<sup>2</sup>. For example, the telephone would be seen as a less interpersonal medium than face-to-face communication as the interpersonal functions carried out by nonverbal behaviours have been 'filtered' out. One of the earliest studies to examine this idea was conducted by Ryan (1976). Participants were asked to solve problems using one of three media to communicate (face-to-face, audio, and video conferencing). From self-ratings conducted by the participants after the tasks, it was observed that users rated the narrower-bandwidth media negatively and far less aesthetically pleasing. In an early book on computer-conferencing, Johansen, Vallee, and Spangler (1979) speculate as to some of the characteristics of CMC group-interaction that could result from filtered-social cues. These characteristics stem from the occlusion of nonverbal indicators and the imperceptibility of cultural, organisational, and motivational cues to behaviour. These characteristics could result in stereotyping, group polarisation, low levels of trust, and groupthink. Conversely, the removal of extraneous nonverbal cues would result in a more productive task-focused interaction, and the

---

<sup>2</sup> What is not always made clear in several literature reviews is that the filtered-cues approach did not preclude the fostering of intimate relationships within CMC. It was occasionally observed that users could adapt to the 'impersonal medium' such that their socio-emotional and relational needs were met.

removal of social protocols might actually promote intimacy. Zimmerman (1987) found that CMC discussions could even have therapeutic benefits, promoting communication between emotionally disturbed adolescents. Not all the effects of filtered social-cues need be negative.

Hiltz and Turoff (1978) describe some of the other effects that can result from the occlusion of nonverbal information. In a pilot study using EIES, the investigators reported difficulties in behaviour coordination that possibly resulted from the absence of conversational turn taking cues. They describe the medium as “much less intimate and self-exposing” whereby users are more likely to feel detached and are less likely to find their social and emotional needs being met. However, Hiltz and Turoff also describe evidence of highly intimate relationships being built up. Users feel the need to compensate for the reduced social cues by making their affective states more explicitly known. In a subsequent pilot study (Hiltz, 1978), it was found that a group of EIES users did not actually view the media as impersonal or difficult. Similarly, in an analysis of the PLANET computer-conferencing system, Ferguson (1977) also found that despite initial discomfort, people soon exhibited enthusiasm for the medium once they had become accustomed to it. Subsequent qualitative work in CMC would then increasingly illustrate the interpersonal side of the medium, but the ‘impersonal’ filtered-cues approach still dominated experimental research throughout the 1980s, most notably through the efforts of the researchers at the Carnegie-Mellon College of Humanities and Social Sciences (Sproull and Kiesler, 1986; Kiesler et al., 1984; McGuire et al., 1987).

In a seminal article, Sara Kiesler and her colleagues (1984) proposed a framework for studying CMC. Based loosely on filtered-cues ideas, the key influences of CMC involved the following six attributes.

1. *The temporal and information-processing pressures of instant messaging*
2. *The absence of regulatory feedback*
3. *The dramaturgical weakness of the medium*
4. *The reduction of status and position cues*
5. *Social anonymity*
6. *Immature etiquette and the norms of CMC culture*

The authors report three experiments comparing group task performance across four settings including (1) face-to-face, (2) email, (3) anonymous computer-conferencing, and (4) non-anonymous computer-conferencing. They found that the CMC groups exhibited the same level of task-focus as the face-to-face groups (not a higher level as was speculated by other authors). They found that dominance by a single team member was less pronounced in the CMC groups and that behaviour was less inhibited. The authors speculated that these results could have arisen due to one of three factors; coordination difficulties, the absence of social influence cues, and depersonalisation.

Outside of the laboratory, Sproull and Kiesler (1986) investigated further the effects of reduced social-context cues in the emails of a large office-equipment firm. Ninety-six employees agreed to maintain copies of all their emails over an eight-week period, and this data served as the

corpora for a content analysis. In addition, the employees received a questionnaire and were interviewed concerning their behaviour across different media. In agreement with the studies reported in the 1984 article, it was found that “electronic mail reduced social context cues, provided information that was relatively self-absorbed, undifferentiated by status, uninhibited, and provided new information” (Sproull and Kiesler, 1986). They also found that people preferred the email medium for sending ‘bad news’. However, once again positive effects were found, including the existence of non-work related discourse that acted to heighten sociability.

### 4.2.3 The SIDE Model

The Social-Identification and Deindividuation Model (SIDE) was developed as part of a collaboration between Martin Lea and Russell Spears (Spears and Lea, 1992). The authors had reviewed the work by the Carnegie-Mellon team under the Filtered-Cues paradigm, and had been especially interested in the observations that concerned the polarisation of online groups. They found the existing explanations for such effects to be problematic (Lea and Spears, 1991) and subsequently embarked upon a new explanation that evoked a contra-classical notion of deindividuation, coupled with a grounding in Self-Categorisation theory. They used this model as a basis for explaining processes of social influence within computer-mediated groups.

Classical Deindividuation theory examines how a person’s sense of individual identity is lowered within a social group. Deindividuated behaviour, in the classical sense, embodies those behaviours that are otherwise inhibited (Zimbardo, 1969) or that reflect the absence of a personal code of conduct. Such behaviour tends to be determined through external, environmental stimuli (Diener, 1979, 1980). Reicher (1984) argued that this approach adopted a classical dichotomous conceptualisation of the self as being in either an operative or an obscured state. In contrast to this model, Reicher adopted a theory of self based upon Self-Categorisation theory (Turner, 1987) in order to gain a new perspective of deindividuation.

Self-Categorisation theory centers upon the ‘social self-concept’ (Turner, 1987). It postulates a hierarchical classification of how the self is envisaged, a system with at least three layers of abstraction, the superordinate, the intermediate, and the subordinate. These correspond to a person’s ‘human’, ‘social’, and ‘personal’ identities respectively. None of these layers has “any privileged status” in defining the self, and the salience of a self-concept in any one of these layers is situation specific. The various factors that contribute to enhancing the salience of a self-concept at one particular level, often act to reduce the salience of self-concepts that exist at the other levels. This creates a “functional antagonism” between the different levels of abstraction. From such assumptions, Turner hypothesises that “there tends to be an inverse relationship between the salience of the personal and social levels of self-categorisation” (p.49) and that the depersonalisation of self at the subordinate level is the driving force behind group phenomena.

Reicher (1984; Reicher and Levine, 1994) theorised that under deindividuated conditions the degree to which an in-group member would conform to group norms depended upon the

demarcation salience between the in-group and the out-group. Reicher attempted to operationalise deindividuation as visual anonymity in a study that tested this hypothesis. In this experiment, participants were anonymised by dressing them up in loose fitting clothes and colour-coded facemasks. The results were disappointing and it was suggested that this might have been due to the poor construct validity of this anonymity manipulation.

Visual anonymity is inherent to the CMC ecosphere and therefore the medium provides a more suitable vehicle for operationalising deindividuation. Spears et al. (1990) recognised this affordance, and subsequently reappraised Reicher's original ideas within a CMC-based scenario. Participants were engaged in group discussions through email and were not brought into visual contact with each other. Pseudonyms were numerical, and were prefixed by the phrase 'group member' or 'participant' to raise either social or personal salience depending upon the experimental condition. Attitudinal questionnaires provided the dependent measure for the study. Deindividuation seemed to result in greater group polarisation when *group* identity was made salient, but less polarisation when *personal* identity was made salient. This experiment seemed to confirm the ideas behind Reicher's Deindividuation theory, but could also provide a framework through which CMC might be understood as a medium. It was this idea that fueled further research into the SIDE model and its application as a theory of computer-mediated communication.

The investigators found difficulties with the explanations forwarded by the Carnegie-Mellon team in describing the group polarisation effects obtained in their experiments (Lea and Spears, 1991; Spears and Lea, 1992). They refuted the filtered-cues hypothesis by demonstrating the social information that was carried within textual paralanguage and in keeping with their 1990 findings, rejected classical deindividuation explanations for polarisation effects (Lea and Spears, 1991). Lea and Spears have advocated the SIDE model as representing the driving force behind social group interaction using CMC (Spears and Lea, 1992). A schematic of the model has been provided in figure 4.1.

CMC can be characterised as having both visual anonymity and physical dislocation, which the authors conceptualise as deindividuating conditions. The affect of this deindividuation is then determined by whether or not *personal* or *social* identity has been made salient. In the absence of any other factors, deindividuation makes people conform to the norms of the salient identity. Therefore, within a chat room where the users do not feel themselves to be a part of a community, there will exist the more personally-oriented and less socially-oriented behaviour of the type observed by the Carnegie-Mellon team. Contrastingly, in a chat room where group identity is made salient then the reverse effect will occur.

Recent studies have examined the predictions posited by SIDE (Lea et al., 2001; Douglas and McGarty, 2001; Postmes et al., 2001). Fifty-six participants took part in a study to test the mediating effect of identity salience on self-categorisation (Lea et al., 2001). Each participant joined a discussion with two other students (confederates of the experimenter) and then filled in a questionnaire. Discussions were held through a text-based CMC program that either was

(identifiable condition) or was not (anonymous condition) augmented with a silent video transmission of the other students. Local group salience was operationalised by assigning a group and group-member name to each participant. Nationality salience was enhanced by telling the participants that the confederates were either British or German nationals. It was found that visual anonymity increased the self-categorisation at the group level (and subsequently group attraction), but would not do the same at the national level. The authors explain that visual anonymity acts to promote the construction of "relatively transient identifications rather than by reinforcing preexisting identifications at the level of wider social categories" (p.536).

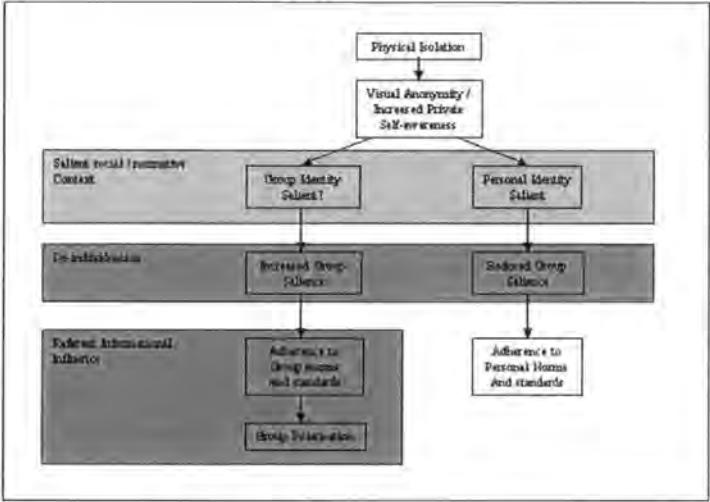


Figure 4.1: The SIDE Model (adapted from Spears and Lea, 1992; p.53)

A more explicit testing of the social influence processes underlying SIDE was reported in the same year (Postmes et al., 2001). Seventy-five undergraduates were split into groups of three before undertaking three tasks. In the first task, the participants were primed with either 'efficiency' or 'pro-social' behaviour using a scrambled-sentence exercise. In the second task, group salience was enhanced with a guessing game with the other group members. This game was manipulated to promote a sense of group cohesion. Finally, the participants took part in group discussions in either an anonymous condition (using text-based CMC) or an identifiable condition (using text-based CMC augmented with pictures of the other group members). An analysis of the discussions showed that participants reflected the primed attitudes in the anonymous condition but not in the identifiability condition. A subsequent experiment was undertaken whereby a neutrally primed subgroup of two people underwent a discussion with a primed subgroup of two, again through either anonymous or identifiable conditions. The neutral subgroup showed conforming behaviour in the anonymous condition, illustrating more directly the social influencing effects of the SIDE model.

The SIDE model is a useful way to contextualise the findings found under the Filtered-Cues paradigm. Nonetheless, this research, like that of previous research was mainly undertaken

using participants who had no prior knowledge of one another. The deindividuation effects posited by SIDE might be subordinate to higher-level relational processes that are probably of more importance when understanding naturalistic CMC interaction. This factor was addressed by Joseph Walther in his formulation of the Social Information Processing approach.

#### **4.2.4 The Social Information Processing Paradigm**

The Social Information Processing (SIP) approach was developed by Joseph Walther (1992) in order to address an apparent incongruity of CMC research. He observed that early experimental studies conflicted with reports from field research by characterising CMC as an impersonal, task-oriented medium. Field research had contrastingly found examples of highly interpersonal CMC interaction, and Walther proposed that this conflict could be accounted for by a more *relational* perspective.

The SIP approach is founded on group of basic assumptions about human behaviour, and an inferred list of propositions. These have been reproduced below (adapted from Walther, 1992):

##### ***I. Assumptions***

- 1. Humans affiliate. They use communication to affect the ways they affiliate, and these messages constitute relational communication.*
- 2. The development of an interpersonal impression of another person is based on the information one gets via nonverbal or verbal-textual channels over the course of several interactions.*
- 3. Developmental change in relational communication will depend on forming an interpersonal impression of another interactant.*
- 4. Relational messages are transmitted (i.e. encoded and decoded) by nonverbal or verbal, linguistic, and textual manipulations.*
- 5. In computer-mediated communication (CMC), messages take longer to process than do those sent face-to-face.*

##### ***II. Propositions***

- 1. Based on Assumptions 2 and 5, the development of interpersonal impressions among previously unacquainted interactants requires more time in CMC than in face-to-face interactions, because CMC takes longer to exchange.*
- 2. Based on Assumptions 2 and 5, personalised communication (based on interpersonal knowledge of others) takes longer to emerge in CMC than in face-to-face interactions.*
- 3. Based on Assumptions 3 and 4, relational communication changes as the number of exchanges increases.*
- 4. Based on Assumptions 3 and 5 and Proposition 1, relational communication in initial interactions is different from that in later interactions.*



5. *Changes in relational communication will take longer to accrue in CMC than in face-to-face interactions.*
6. *Based on Assumptions 1 through 5, given sufficient time and message exchanges for interpersonal impression formation and relational development to accrue, and all other things being equal, relational valences in later periods of CMC and face-to-face communication will be the same.*

The central thesis of the approach is that CMC users are social actors that can adapt whatever interpersonal cues are made available to them for the purpose of relational management. This process may take longer in a medium such as CMC, where there is reduced transmission speed and parallelism than there is face-to-face, however it will still occur given enough time. To paraphrase Walther (1992), the process itself is driven by relational motivators that prompt users to form impressions by decoding text-based cues, deriving psychological knowledge about their collocutors. Thus, they manage relational changes and encode relational messages. In terms of impression formation, Walther refers to ideas that underlie the development of an interpersonal epistemology. CMC users will use a variety of text-based probes in order to ascertain the relational function normally embodied through nonverbal behaviour.

Walther (1996) states that although field studies had refuted earlier experimental work, they had failed to account for the impersonality findings, "in exorcising the external validity of the work, they have ignored the question of internal validity, throwing the empirical baby out with the theoretical bathwater and ignoring lessons that might yet be instructive" (p.9). The results obtained by the early experimental studies can be explained by SIP.

The early experiments were often conducted with zero-history groups over a short period. In addition, the CMC conditions were often compared to face-to-face conditions over an equivalent period. Walther (1992) suggests that in these experiments impersonality had sufficient time to develop in the face-to-face conditions but not in the CMC conditions. The results were being overgeneralised as a characteristic of the medium, but were in fact simply a function of the experimental conditions. Understanding these conditions of impersonality is relevant because in some scenarios (such as in group decision support systems), such strategic conditions may be of use in the democratisation of interaction.

Walther and his colleagues have subsequently conducted empirical studies that support the SIP approach. In one study, Walther (1993) partially demonstrated the longitudinal effects of relational development in CMC. CMC and face-to-face conditions were each comprised of sixteen, zero-history, three-member groups that were set to work on three sequential tasks over a number of weeks. Intra-group contact was made solely within asynchronous CMC or within a private classroom (depending on the condition). Relationship development was measured upon completion of each task. In both conditions, impression development occurred over time. Although the levels in CMC did not quite manage to 'catch up' with the levels in the face-to-face condition, Walther suggested that this would have happened given longer time. Other studies have also found that the temporal dimensions of CMC communication (such as the frequency

and duration of postings) affect impression development (Liu et al., 2001). Walther and Tidwell (1995) examined the role of 'chronemics' (short-term temporal factors) within the context of email. Participants using asynchronous CMC might assess temporality along such cues as message timestamps, the interval between inbox checks, ordinal threading, or real-time notification (Walther and Tidwell, 1995). In their study, Walther and Tidwell (1995) examined the perceptual relevance of message timestamps. Experienced CMC users were asked to rate reprints of email dyads along scales of intimacy and dominance. The two main types of dyad were comprised of either task-oriented or socially-oriented discourse, posted during task-oriented or socially-oriented times of the day (relating to work hours), and at different frequencies between posts. Some hypotheses were validated (such as greater dominance being attributed to task-oriented emails sent outside of business hours), however a significant interaction effect led the authors to believe that the relationship between the factors investigated were more complex than was initially assumed.

In another study, Walther (1994) identified that anticipated future interaction in CMC was also an important factor in relationship development, more so than it was face-to-face. This highlighted the significance of an implicitness in the SIP approach (the nature of participant motivation). Utz (2000) examined this factor in a study that focused on the specific communicative processes (as opposed to simply the chronemics) by which impression development took place. Utz surveyed one MUD community and found, somewhat expectedly, that skepticism towards MUDs for the purpose of relational development affected that development. She also demonstrated the effectiveness of emoticons for transmitting relational information. The interpersonal power of textual paralanguage has also been reported elsewhere. It has been surmised, in keeping with the filtered-cues approach, that flaming behaviour was often the product of the misinterpretation of postings resulting from the occlusion of nonverbal cues (Franco et al., 1995; Spitzer, 1986). Thompson and Foulgar (1996) found that the use of pictographs actually helped to moderate the perception of flames in email.

SIP can account for the impersonal and the interpersonal relationships demonstrated in previous studies. However, some field studies go further to suggest that in some cases, CMC interaction can actually exceed the intimacy that is gained through comparable face-to-face interaction. Walther (1996) addressed the case of *hyperpersonal* interaction in a later article. Walther proposes four means by which a researcher might examine hyperpersonal interaction, through the receiver, the sender, channel characteristics, and feedback.

Walther (1996) suggests that under certain circumstances, the "receivers inflate the perceptions they form about their partners" (p.17). Walther compares this observation with the polarisation findings revealed by the SIDE model. In combination with this impression formation process, the sender is engaged in a process of highly selective impression management. This heightened impression management is afforded by the reduced cues and the channel characteristics of asynchronous CMC. These characteristics allow the sender to divert cognitive resource into message construction and away from nonverbal presentation. Asynchrony can also alleviate

temporal response pressure and turn taking competition. Walther refers to McGrath's concept of *entrainment* by which collocutors negotiate "conflicting temporal interests and requirements" (McGrath, 1991, p.162). Asynchrony mediates the conflict by providing people with as much time as they need. As the conflict is mediated, Walther observes that this also mediates the task-impeding affect of socially-oriented exchanges in group interaction. The whole process of heightened impression management and formation is then reciprocally fed back between collocutors, confirming impressions and militating the hyperpersonal effect. Therefore, it turns out that heightened levels of intimacy are prevalent within CMC *because* of its interpersonal nature. Hardey (2002) found that respondents using an Internet dating service managed to instigate relationships that could not have occurred offline. However, some of these relationships were not successfully brought offline due to those difficulties seemingly overcome by CMC (such as geographical separation). Inflated impressions of a person's appearance were also reported along with disappointment when the participants did eventually meet face-to-face. It would seem that hyperpersonality in many cases can lead to negative as well as positive effects.

The SIP approach provides a means of studying CMC in a way that is compatible with the body of empirical CMC findings. Walther (1996) mentions that in future work the intentionality of the user must be examined. For example, what happens when a user wishes to be malevolent or deceitful? Additionally, it is surmised that the approach may be somewhat elitist in the demographic sample to which it pertains. It is a task for future research to see to what extent this really is the case.

#### **4.2.5 An Ecological Approach**

James Gibson's controversial book, 'The Ecological Approach to Visual Perception' was published in 1979. This was to become the last book written by Gibson and it represented a culmination of his life-long work on perception theory. Many of Gibson's previous ideas were reformulated in a treatise that challenged the prevalent approaches to the study of perception at that time. More relevantly, the heart of this treatise has promulgated to other areas of psychology and has recently been proposed as an investigative framework for the study of communicative technologies (Gibson, 1979; Hutchby, 2001).

Hutchby (2001) identifies that much of the recent work pervading the 'sociology of technology' has attempted to eschew technological determinism by adopting some form of social constructivist position. Such positions fall either into the interactionist or into the socio-technical network approach. The focus of the interactionist approach centers upon the interplay of the technical and the social. The network approaches however, aim to break down the old analytic dichotomy between these elements, instead focusing on a network of actors that may be either social or technological. Hutchby cites Grint and Woolgar's (1997) critique that these approaches still contain a "residual technicism" whereby artifacts may exhibit qualities that are free from subjective interpretation. However, Hutchby rejects Grint and Woolgar's reversion to an even

more radical constructivism, suggesting that such subjectivisation takes the analytical locus away from the everyday common-sense knowledge that any one artifact may differ in the set of possible interpretations that it may support distinctly from another. This theoretic appraisal is much like Gibson's own appraisal of the state of the art in visual perception research. Gibson's concern was in removing the old dichotomy between the perceiver and the perceived without reverting to radically constructivist cognitivism. Like Gibson, Hutchby suggests a more functional perspective. This perspective focuses on the affordances of the actors and technology, a stance grounded in an unconventional subjecto-objectivism.

Any object has a certain set of 'opportunities for action' that it can afford a particular organism. In Gibson's words, "the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (Gibson, 1979, p.127). For example, a flat brick wall may afford a spider 'climb-on-ability' or may afford a human 'lean-on-ability'. The affordance refers to both the environment and the organism, and so the affordance of 'climb-on-ability' for a spider on a brick wall may not necessarily transfer to another organism, such as a human. Affordances represent a relation between the organism and the environment such that they cannot be described as distinct from those two elements. Thus, from a human perspective what is of interest is the discovery (and the construction) of a world of artifacts revealing affordances that specifically relate to human interests.

Hutchby (2001) suggests that researchers should focus on the affordances of the mediating technology on the user, and to ask what opportunities for communicative action are relevant for that technology. For example, in CMC a researcher may wish to examine what affordances for communicative action exist within the textual medium or within the user interface itself. Hutchby is not the only author to suggest this application. Norman (1988) has also adapted affordance theory in his book 'The Psychology of Everyday Things'. However, Norman's adaptation seems to dispense with the relational mutualism of affordance theory, by making the affordance an attribute of the object alone. It seems unclear as to what benefit this serves, as it draws attention away from the relevance of the object to a particular organism. It is clear from Gibson's writings that this is precisely the interpretation that Gibson wished to discourage.

Another approach to affordance theory in CMC can be derived from recent writings on the 'socialisation of affordances'. Instead of looking at the affordances of the interface, a researcher might draw their attention to the affordances for *social* action between two organisms. This idea was suggested by Gibson, and later taken up in two seminal articles calling for a socialisation of the affordance approach (McArthur and Baron, 1983; Knowles and Smith 1982). So in face-to-face interaction, the researcher may attune to what meaning is perceived from social events; meaning that is communicated through nonverbal behaviours. This approach to impression formation can be similarly applied to the paralinguistic aspects of CMC. For example, does the user's posting speed (a paralinguistic cue) afford a certain type of social action for one type of user over another? The inchoate socialisation of Gibsonian theory is still debated in the literature, and has yet to be applied to CMC. However, it is an analytically relevant approach as

it focuses on the 'what' of social perception rather than the 'how' and this is a central distinction between the face-to-face and CMC domains.

Finally, it should be noted that one of the central critiques to the affordance approach lies in the processes of attunement and culturisation (Schmitt, 1987). The approach says little of how one organism selects a particular affordance over another, and how this may be affected by cultural differences. It seems as though it may not be possible to describe the socially negotiated value and meaning of the world solely in terms of physical stimuli. However, it can also be said that the approach does not need to set such all encompassing goals in order to be of analytical use and has complementary value to other theories such as SIP and SIDE.

### 4.3 The Structural Features of Quasi-synchronous CMC

A central question addressed in the field of CMC is whether textual discourse resembles other written or spoken forms of mediated communication, and whether this type of communication could eventually change the ways people use language both online and offline (Baron, 1984). This question has led to the study of the structural features of CMC using applied conversation-analysis techniques that focus on the relational characteristics of textual discourse, particularly with regard to chat-room interaction.

The style of discourse enacted within a chat room system like IRC is often referred to as being either synchronous or quasi-synchronous (Garcia and Jacobs, 1999). Synchronous *written* communication is a novel form of interaction, whereby collocutors need to be communicating simultaneously in order to facilitate successful discourse. Quasi-synchronicity recognises the affordance of many chat programs to allow utterances to be edited and revised before they are posted. Other synchronous, real-time styles of communication, such as natural conversation, depend upon a wealth of nonverbal and paralinguistic cues in order to attain successful co-ordination. Early advocates of the filtered-cues hypothesis believed that the absence of these co-ordination mechanisms would lead to the break down of successful quasi-synchronous conversations. However, theories such as SIP predict that people are able to utilise the affordances of the medium to successfully manage their discourse.

Central to successful conversation are the mechanisms underlying turn-taking and speaker roles (Sacks et al., 1974). In an article by Garcia and Jacobs (1999) the CMC environment is described as having a 'temporal multidimensionality' that is not present in face-to-face conversation. Collocutors do not share the passage of conversation in the same time frame; the speaker experiences the construction of the utterance before the hearer, and in fact, the listeners have no access to the utterance construction process at all. This feature is responsible for the reinterpretation of the transition-relevance place (TRP), the decomposition of conventional participant 'roles', and a nullification of conventional turn competition (Garcia and Jacobs, 1999; Hutchby, 2001; Herring, 1999).

The TRP no longer deals with a transition of speaker, but represents an opportunity for any collocutor to begin constructing an utterance (Garcia and Jacobs, 1999). This means that there

is no clear 'next speaker' until that person has posted, and this provides multiple participants with the opportunity to self-select. This can result in the production of 'phantom adjacency pairs' that occur when two utterances are produced independently in response to previous utterance, but are posted at such a time that they appear like an adjacency pair themselves. For example, one speaker may be typing an answer to a previous question, whereby a second speaker is posing a new, different question. However when these two utterances are posted, the first speaker's 'answer' looks like a response to the second speaker's question, simply because the second speaker had managed to post first.

Experienced CMC users have found ways to adapt to the peculiarities of the TRP. Conventions such as 'one line, one turn' or post-fixing a '%' to indicate that a current speaker has yet to continue, are just some of the ways in which people make it clear about the transition of speakership (Herring, 1999). Other methods include explicitly naming the next speaker, referred to as 'addressivity' (Herring, 1999) or 'nominalisation' (Hutchby, 2001).

Herring (1989) points out that the skewed form of turn taking that results from quasi synchronicity and the reinterpretation of the TRP, is in fact a more efficient form of turn taking. In the face-to-face scenario, it takes just as long to hear an utterance than it does to speak an utterance. However, this is not the case in CMC, where it is quicker to read than it is to type. Subsequently a number of people can be constructing their responses to an utterance at the same time (as opposed to sequentially), and this is an efficient parallelism that cannot be replicated offline.

Garcia and Jacobs (1999) identify the following roles of collocutors in a QS-CMC conversation. It is possible for any collocutor to actually be performing multiple roles at once.

- *Message Constructor - Participant is typing an utterance*
- *Message Poster - Participant is sending the utterance to the other participants*
- *Waiter - Participant is awaiting the receipt of an utterance*
- *Reader - Participant is reading an utterance*
- *Worker - Participant is performing some other task during a CMC session*

Garcia and Jacobs discuss the turn taking options that occur at the TRP. When the current speaker wishes to select another speaker to talk, there may be problems. The recipient may not be 'waiting' and instead may be reading, working, or constructing (in other words not 'listening'). Alternatively, the response may be interrupted by another speaker who has decided to self-select. These phenomena echo what was said earlier about quasi synchronicity, a reader "cannot rely on sequential organisation to locate referents of utterances" (Garcia and Jacobs 1999, p.356). It is not possible for the current speaker to select the next poster, merely a future poster.

Drawing from studies that examined telephonic communication, Rintel and Pittam (1997) explored opening sequences in IRC. They derived a general opening sequence that described a user's introduction to an IRC session, starting with an automated 'announcement' by the IRC

program that acts in a similar manner to the telephone but with different affordances for action. The first part of this sequence, the Channel Entry Phase (CEP) was elaborated in a further paper and this has been outline below (Rintel, et al., 2001).

1. *Automated Joining Event*

*The IRC system broadcasts a message indicating the entry of a new user (the Joining Announcement), and similarly confirms to that new user that they are now able to beginning chatting (the Joining Confirmation).*

2. *Joining Initial Behaviour*

New user posts either a collectively or individually addressed utterance. Simultaneously, another user might post a Joining Initial Reaction.

3. *Responses to either the Joining Initial Behaviour or the Joining Initial Reaction.*

Each stage presents a number of different affordances for action and so there are several different paths through the sequence. The authors observe that there seems to be a considerable amount of ambiguity in the CEP concerning whether or not the AJE counts as an initial utterance by the new user. The existing members of the chat will either ignore the AJE or respond to it with a joining confirmation. As the authors remark, "one of the difficulties of ratification is that public IRC channels are forums more conducive to collectivity than dyadicy" (Rintel et al., 2001). Users may attempt to align to one another in these initial stages through considerable paralinguistic matching during greetings. Adornments involving capitalisation and non-alphanumeric characters are used in a similar way to nonverbal matching in face-to-face conversation.

The conversation-analytic approach to the study of QS-CMC is still in its early stages. This type of research provides a great opportunity for understanding the adaptive methods by which people are able to communicate within novel media. One of its key uses will be to explore the nature of *group* conversations, ubiquitous within virtual communities. This area is often eschewed by communication researchers who model most of their theories on *dyadic* communication processes.

## 4.4 The Self Online

The social infrastructures supported by CMC provide new contexts within which users can continue the construction and maintenance of their 'selves'. The nature of the 'self' online has been the focus of a considerable proportion of CMC research and this is unsurprising. The self is in many ways dependant upon the interpersonal context in which it is situated and CMC-based communities represent a significant modal departure from interaction experienced in everyday life.

Probably the most significant feature of a CMC-based context is the greater level of anonymity it affords the user. This in turn, provides the user with great control over their self-presentation to others, and ultimately, through self-fulfilling prophecy leads to a fashioning of their own self-

perception. This has led to a number of postmodern writers postulating that within certain contexts the online personae constitutes a 'society of selves' (Turkle, 1996). Users who engage in role-playing within MUDs or BBSs may have a number of different identities according to the particular MUD they are in, or the particular mode of interaction they wish to encourage. For example, rooms may be created for users to come out of character (Myers, 1987), possibly to allow discourse of the more pragmatic aspects of society regulation to take place. Alternatively, users may adopt one pseudonym for close friends and another for when greater anonymity is sought (Waskull and Douglass, 1997). In a survey of lamdaMOO by Schiano (1997), it was found that about 25% of respondents changed their character within a two-week period, and that they usually maintained about three different identities on average. Despite this, most players spent most of their time using a single main identity. In a survey, conducted by Roberts and Parks (1999) gender switching was reported by 40% of members in social MOOs and 56.7% of members in role-playing MOOs. In agreement with the Schiano survey, most of the members engaged in gender swapping for less than 10% of the time. The purpose of this behaviour was usually reported as some form of exploration of identity, and the reconstruction of the conventional dichotomous conceptualisation of gender.

The concept of a user having multiple selves, both online and offline should be seen as metaphoric. There exists only one single psychological self that is exposed to a variety of integrated experiences (Baumeister, 1998). The entirety of self-knowledge is too great to be in awareness at any given moment, and the part of self-knowledge that is at the forefront of awareness has been referred to as the 'phenomenal self' (Jones and Gerard, 1967). The phenomenal self needs internal consistency, but can contradict that which lies outside of this currently primed knowledge (Baumeister, 1998). This is a useful way to conceptualise a person's shifting identity within the CMC domain.

Self-presentation may be specific to the adopted identity in both content and form. One way in which the offline self is presented is through the body; a body occluded in the CMC domain. Users may choose whether to disclose their physical characteristics, lie about them, omit them completely, or reinvent them online. For some people, stigmatised because of their appearance or disability, this may represent an opportunity to explore the less salient facets of their identity with other people (Bowker and Tuffin, 2002). Bower and Tuffin (2002) report the case of a disabled lady who reported that her online interactions represented a 'mental break' from her impairment. In this case, reality was shifted away from a visual ontology and towards a linguistic one. The body may also be constructed online as has often been found in MUDs. This can range from the verbal description of certain bodily actions such as handshaking (Argyle and Shields, 1996) to constructions that are more elaborate. McRae (1997) reports one community where the users are represented by different types of animal in a social structure dictated by species. At first, it seems natural to distinguish such environments as simple role-play, but unlike role-play, there is no natural terminus. The interaction is ongoing and this can increase the power and significance of this ontology (Turkle, 1996). The online self may not exist in the same way (if at all) than the offline self does. Nonetheless, the social-constructivist perspective



has been put forward that a person's *belief* in such a separate online self is what is important for researchers studying online communities (Waskull and Douglass, 1997). The body online has also been described by postmodernists as prosthetic to the offline body. The so-called cyborg identity (Haraway, 1991) defines a certain symbiotic, augmentative relationship between the user and their technological representation. This relationship has also been described as a form of transcendence of identity rather than a simple prosthesis (Wilbur, 1997). Waskull and Douglass (1997) use the term 'cyberself' to refer to the online identity, "a cyberself is the meaning of personhood (or experience of personal identity) emergent within dislocated and disembodied forms of electronic-computer-mediated interaction between persons and/or communication technologies" (p.386). The key parts of the definition are dislocation and disembodiment. Within the unitary approach to self, the cyberself actually represents that phenomenal self activated within a particular online context.

Self-presentation is also evident through pseudonymity. Pseudonyms are often (but not always) chosen in order to make some sort of statement, explicitly hiding identity whilst revealing some other aspect of personality (Jaffe et al., 1999). A content analysis by Waskull and Douglass (1997) categorised pseudonyms into six main families; real-name variations, geographical references, interests, occupations, lifestyles, and motives. Pseudonyms provide a very efficient way of evoking a first impression, through use of stereotypical and culturally defined markers. The most common of these is that of gender, and gender deception is prevalent within romantic CMC chat-rooms (McRae, 1997). In some places, such as LamdaMOO (a long established virtual community) gender has been redefined altogether. In LamdaMOO, neutral gender pronouns are in frequent use, but genders are also far more configurable representing a freedom from the basic male-female dichotomy (McRae, 1997).

Through the manipulation of the online identity, a user is explicitly engaged in a process of reflective image construction. This might involve the creation of more powerful and less inhibited identities that can provide a person with more perspicacious self-knowledge and self-exploration (Bromberg, 1996). This might or might not have therapeutic applications. Turkle (1996) describes two case studies exemplifying how interaction within a MUD can allow people to either work-through their problems, or simply act them out within a different domain. One resulting effect can be that users end up placing a higher value on their online identities than they do on their offline identities (Markham, 1998).

In an analysis of the self-ratings of fifty-five participants, Matheson and Zanna (1988) found some preliminary evidence that showed that users in a simple conferencing system were more self-aware than other users engaged in face-to-face discussions. This could be a significant finding despite being mediated by weak construct validity and a temporally restricted scenario. People have been found to enjoy self-awareness in situations where they feel that they are superior to others, and self-awareness has been known to encourage behaviour that reflects a person's positive attitudes (Carver and Scheier, 1981). Scheier and Carver (1977) have observed that increased self-awareness can intensify emotional reactions. If CMC truly

heightened self-awareness, then this finding would provide more evidence reflecting Walther's (1996) ideas on hyperpersonality in CMC. Conversely, self-awareness has also been correlated with negative states such as anxiety, alcoholism, schizophrenia and depression (Greenberg and Pyszczynski, 1986; Ingram, 1990).

The CMC medium allows interaction with a vast amount of other people from a variety of communities that transcend those that are closely geographically located. This state of 'social saturation' (Gergen, 1991) can provide a means by which those people with identities marginalised within their local community can find like-minded others. McKenna and Bargh (1998) undertook a content analysis of both marginalised and mainstream newsgroup communities. They found some evidence to suggest that there was greater participant involvement in those newsgroups representing marginalised identities (identities 'concealable' in everyday life). This involvement may give rise to greater levels of self-esteem for the otherwise marginalised self, possibly encouraging a person to reveal that aspect of their identities within their offline community.

Social roles define another important interpersonal aspect of self. Online communities present many opportunities to adopt a particular role within society. Examples in MUDs can include wizards, system administrators, 'vigilantes', 'murderers', and 'rapists' (Ito, 1997; Dibbel, 1996). Myers (1987) identified three types of communication leader within a BBS set up over two months; system operators, system experts, and social experts. Virtual weddings are increasingly common as well as other types of familial relationships (Turtle, 1996; Ito, 1985). The extent to which such roles affect the self has yet to be explored in any great depth, but are increasingly relevant when it comes to controversies concerning 'virtual infidelity' and 'virtual statutory rape' (Waskull and Douglass, 1997).

## **4.5 Deception in CMC**

The anonymous and dislocated features of CMC provide users with various opportunities for deception. For example, within chat rooms it is almost 'common wisdom' that most women (especially promiscuous ones) are most likely to be men (Curtis, 1992). Sometimes this form of deceit may be benign or playful, but sometimes it is not. This section will examine some cases of Internet deception, and will review some very recent attempts to understand the phenomenon. This will lay the foundation for the following chapter, which will attempt to synergise the research on CMC and deception.

### **4.5.1 Case Studies**

One of the classic readings of Internet deception is Lindsay Van Gelder's (1996) account of 'The Strange Case of the Electronic Lover'. The story concerned a prominent psychiatrist, in his early fifties, that assumed the identity of a fictitious character *Joan* within one of the early CompuServe channels. Joan was a vivacious lively character who was a complement to the personality of the psychiatrist (referred to as Alex). The character Joan, was a victim of a horrific

car accident that had left her disfigured, crippled, and without speech. Alex had found, almost by accident that by adopting a female personae, he was able get other women to 'open up' to him, but the deceit extended far beyond the realms of professionalism. Joan engaged in numerous supportive and intimate relationships with other people, all of who were unaware of the masquerade. Alex admitted the deception after certain doubts were raised by the other channel members. People had observed contradictions in Joan's statements but most fundamental of all was Joan's inability to demonstrate an offline existence. This manipulation was successful for so long because of Joan's apparent immobility and the shame of her appearance. After the deceit was uncovered, many members continued their relationships with the newly emerged Alex (originally introduced to the group through 'Joan'). Many believed that within CMC, physicality was not an issue and that the soul of Joan and Alex were inseparable. This story, illustrates the diversity of feelings toward online deceit, the ways in which deceivers can be uncovered, and the means by which people may find themselves trapped and seduced into an ever-thickening mask of lies.

Perhaps the most pressing, if not salient concern of chat room deception, is the potential for paedophiles to utilise the medium. There are an increasing number of reports that detail how paedophiles have lured and groomed victims via the mask of CMC. Children are also at risk from other adolescents as well. In a report published by the Home Office, it was thought that up to a third of all sex crimes are perpetrated by adolescents [Home Office, 1998]. A number of initiatives have been organised to better inform children, parents, and teachers of the dangers that prevail within chat rooms. Examples of these initiatives include the 'Chat Safe' programme in the UK (Thames Valley MET and the RAF), 'Innocent Images' in the US (FBI), and Chatdanger.com (Childnet International).

Quayle and Taylor (2001), report a case study of man charged for downloading child pornography from the Internet. The man (Q.X) did not initially intend to groom and lure children over the Internet, but this progression occurred as he was increasingly drawn into the medium. Initially Q.X adopted a false persona, Joe, over a twelve-month period. The description of Joe's appearance was taken from what Q.X would have looked like when he was eleven. Q.X believes he was able to enhance his 'role-play' because of his interest in creative writing and his vocation as a schoolteacher. Initially, Q.X. was seeking a form of 'romantic companionship'; a motivation that echo's the FAQs that can be found on paedophile newsgroups. Interestingly, a couple of the young boys that Q.X. did start relationships with online turned out to be other paedophiles, and hypocritically this caused great offence to Q.X.. Initially Q.X. did not like associating with chat room members who were very blunt, demanding cybersex and offline meetings, but this was precisely the behaviour that Q.X started to adopt during a short period leading up to his arrest.

There have been a number of incidences of other paedophile convictions reported in the press. Patrick Green was sentenced to five years in prison after having sexually abused a thirteen-year-old girl that he met in an Internet chat room [BBC 24/10/2000]. A BBC correspondent

reports talking to 'Georgie', a thirteen year old girl who befriended a fifteen year old boy over the internet. This boy turned out to be a forty-seven year old man who was arrested by police, but not convicted (at that time under UK law, no crime had yet been committed) [BBC 15/03/2001]. Katherine Tarbox (2001) relates her story in the book 'Katie.com' where she went to meet 'Mark' a forty-one year old man who had masqueraded as a twenty-three year old. He sexually assaulted her and was eventually convicted under the US Communications Decency Act. In the US, detective Rick White claimed to have encountered a potential pedophile within seconds of having entered a chat room. Detective White had masqueraded as a fifteen-year-old boy, and had found a 'twenty-seven year old women' who expressed desires to engage in cybersex, and who expressed a fondness for pictures of eight-year-old girls [CNN 24/07/2000].

The last example illustrates the other side of the coin for deception on the Internet. Identity deception can also be used by law enforcement officers to entrap paedophiles. ZDNet, the online technology news service, set up an investigation of Internet pedophiles, by going into certain Yahoo messenger chat rooms under the alias 'Tina\_Bell' and recording the resulting conversations. Their investigation showed examples of phone sex offers, sexual role-play, and cybersex [ZDNET 20/11/2002].

#### **4.5.2 Recent Research**

Only in the past five years, has there been any significant enquiry into the problems of online deceit. Some early survey evidence has suggested that there are gender differences in deceitful behaviour online. It was found that men were more likely to lie in order to feel more open, whereas women and younger people were more likely to lie for the sake of safety (Whitty, 2002; Whitty and Gavin, 2001). In another survey, it was found that people engaging in online relationships were more likely to misrepresent themselves than people involved in face-to-face relationships would (Cornwell and Lundgren, 2001). This correlated with the attitudes of the respondents who saw their relationships as less serious and less committal. Such results are tentative at best and should at least be constrained to the particular type of relationship under analysis. For instance, Turkle (1995) identifies a typology of three classes of relationship, the real-life, the virtual with the real, and the virtual with the virtual. It might be reasonable to conclude that people may not view their virtual-to-virtual relationships as particularly serious, but cases have been reported of relationships where someone who is disclosing themselves fully is involved with someone who is masquerading their identity (Turtle, 1995; Rheingold, 1993; Van Gelder, 1996). The effects that arise out of the revelation of such a masquerade are usually sufficient to suggest that some of the participants do not view their online relationships lightly.

In her study of Usenet, Judith Donath (2001) applied a cost-related framework for understanding online identity deception that is borrowed from the animal kingdom. Within a sender-receiver model of communication, she distinguishes between 'assessment' and 'conventional' signals within the context of Zahavi's Handicap Principle (Zahavi, 1993). An assessment signal is the optimal signal, "that best amplifies the asymmetry between an honest

signaler and a cheater" (Donath, 2001). These signals are closely related to the trait being advertised, and are costly so that they are very difficult to fake. Therefore, an email signature with a link to a personal homepage within an institution is an assessment signal, and this contrasts with conventional signals such as responses to A/S/L requests within a chat-room. Within certain Usenet groups, trust is enhanced through the transmission of assessment signals of identity. Such signals are often sought when seeking advice from so-called 'experts' (one of the main functions of Usenet). Donath provides an example of people demanding to see pictures of a man claiming to be an expert bodybuilder when a difference of opinion arose. Such pictures were refused, subsequently undermining that person's credibility.

As technology changes there is a shift in those signals that are considered as *assessment* and those that are considered as *conventional*. Donath (2001) states that the email 'From' field is an example of a signal that at one time could be considered an assessment signal, but now is considered a conventional signal due to ease at which modern software allows the user to manipulate such information. Donath identifies one form of category deception prevalent in newsgroups. Trolls are people who attempt to cause arguments and flame wars online for their own amusement. According to Donath (2001), "the troll attempts to pass as a legitimate participant sharing the groups common interests and concerns". In many ways, the success of the troll depends on their knowledge of the groups identity cues, and their clever manipulation of conventional signals that may be perceived by other members as more reliable than should be the case.

Another approach to the analysis and detection of CMC deception has very recently been developed by Judee Burgoon and her colleagues (Burgoon et al., 1999). This treatment involves the synthesis of Interpersonal Deception Theory, and to some extent, the relational ideas presented in SIP and the Gibsonian Ecological approach. These ideas are explored and expanded upon in the next chapter.

## 4.6 Summary

This chapter started with a review of the primary theoretical outlooks that have been adopted within CMC research. It was demonstrated that the older characterisations of CMC as an inherently impersonal medium were poorly contextualised and that it is possible for highly intimate relationships to be developed using the medium. In fact, other processes exist that can cause hyperpersonal relationships to develop, and this bias has implications for the level of trust that participants may afford to potentially deceptive collocutors. CMC users have been able to overcome some of the discursive difficulties of coordination by reinterpreting the transition-relevance place and taking advantage of a new constitutive system of speaker roles. Social information can also be expressed via various 'paralinguistic' cues, the importance of which was underplayed by early research. The concept of the self is transformed when taken into the online context and the motivations underlying identity deceit are multiple and not always

intentional. There is a clear motivation for researching the nature of online deceit, as the repercussions of these episodes are usually not confined to the 'virtual' domain.

This literature survey has focused on a variety of material and some correlates can be drawn between that research. Interpersonal Deception Theory as a communication theory has much in common with the contemporary theories of CMC in its attempt to focus on the interpersonal and its consideration of the importance of context. The verbal clues to lying presented in Chapter 3 may also be applied to CMC discourse, although these techniques will have to be adapted to the qualitative and structural peculiarities of this mode of interaction. The affective factors driving the physiological clues to deceit (examined in the previous chapter) may be exhibited in different ways through a user's interaction with their computer. In addition, the sort of identity deception prevalent within CMC provides a novel type of lying in which the relevant clues have yet to be expounded.

In the next chapter, the material explored within this literature review will be synergised and complemented with some pilot research. This treatise will hopefully provide a direction for researchers who wish to tackle the difficulties proposed by CMC deceit and who wish to explore further the possibility of online lie-detection.

## Chapter 5

# Exploring Interpersonal Deceit, Lie-Detection, and CMC

This chapter explores computer-mediated lying and deceit. These themes can be studied from two different perspectives. The first perspective aims to understand the processes and clues to interpersonal deceit within online communities and forums. Lie-detection research that adopts this perspective will use passive techniques of observation and measurement in order to investigate and detect deception. The second perspective uses CMC as a lie-detection tool within controlled scenarios similar to a SVA interview or a polygraph test. This perspective is more active, and affords the investigator a greater level of power in manipulating the measurement software, hardware, and the deceptive context. These two perspectives are closely related and benefit from similar investigations. The focus within this chapter will be on the first perspective, but a great deal of this theorising is relevant to the second perspective as well. The chapter begins with a discussion of lie-detection from these two main approaches and a strategy for naturalistic lie-detection is outlined. The propositions of Interpersonal Deception Theory are applied to the domain and the non-strategic leakage clues, posited by this theory are expounded. This is followed by a discussion of strategic deceptive action. This treatment is supported with qualitative research exploring the types of CMC deceit, the techniques used by deceivers, and the clues that heighten suspicion. A categorisation scheme is presented that suits the concerns of naturalistic lie-detection. After a summary of the chapter, the need for future research is discussed with respect to the two perspectives, and this is related to the research provided throughout the rest of this thesis.

---

### 5.1 Lie-detection and CMC

The evolution of human interpersonal communication has facilitated the development of quite sophisticated deceptive behaviours. As a relatively new form of communication medium, CMC chat systems represent both new problems and new opportunities for lie-detection research. The purpose of this chapter is to explore this next step in the evolution of deceptive interpersonal communication in greater depth.

Lie-detection can be approached from two perspectives, that of altering a technique to suit a deceptive context (naturalistic lie-detection), or that of altering the context to suit a technique (controlled lie-detection). Naturalistic lie-detection is typically unobtrusive and observant, and represents the minority of research because of the intrinsic difficulties discussed in the previous chapters. Controlled lie-detection, such as polygraph tests, SVA procedures, and laboratory experiments represent the most common (and successful) form of research because they allow the investigator to constrain the ways in which the examinee is able to lie. Techniques differ according to the extent that they can be considered either controlled or naturalistic, and most lie along a continuum between these two poles.

Towards the pole of controlled lie-detection, quasi-synchronous CMC (QS-CMC) presents new opportunities to control the deceptive context. In Chapter 3, some of the desirable qualities of an SVA interview were presented including (1) the reduction of anxiety, (2) the minimisation of data contamination, and (3) the promotion of free-recall. A QS-CMC interview may reduce anxiety through dislocation, may minimise corruption through semi-automation, highly controlled questioning, and the removal of nonverbal cues, and may help promote recall through immersion. QS-CMC is also highly interactive and can keep many of the advantages of face-to-face interviewing whilst also eliminating the need for manual transcription. QS-CMC may provide opportunities, complementary or otherwise for existing controlled lie-detection techniques. In addition, the monitoring of computer-based control provides the opportunity to record a variety of low-level data that could reveal useful diagnostic 'nonverbal' information that has hitherto remained unexplored.

In Chapter 4, some examples of the harmful consequences of chat room deceit were discussed, revealing the motivation to explore ways that this form of deceit can be detected or otherwise prevented. There are, of course, places where lying through CMC is either innocuous or indeed prudent, and so it is important for researchers to establish what sort of lying is prevalent online, what is of the most concern, what is the peculiar nature of this form of lying, and what clues can be used to aid in its detection?

So from the perspectives of controlled and naturalistic lie-detection, two broad questions can be raised:

1. What is the nature of online lying, and how could techniques be either created or adapted in order to detect it?
2. How might QS-CMC be utilised in order to provide additional control and measurement within a controlled lie-detection procedure?

The exploration of these two questions forms the broad basis for the research contained within this thesis. The first question will be discussed within this chapter, and the second will be discussed in the three chapters that follow.

This chapter attempts the following:

- *to set out a practical strategy by which naturalistic lie-detection can be undertaken*
- *to explore the mechanics of lying by synthesising computer-mediated theory with Interpersonal Deception Theory*
- *to expound the non-strategic behaviours that could provide clues for naturalistic lie-detection*
- *to describe research that was carried out as part of this thesis, in order to explore the complexity of deceptive strategic behaviour in depth*
- *to describe a categorisation system for classifying different types of lying behaviour performed in CMC*



## 5.2 A Process for a Practitioner

Lying and deceit are behaviours that can be promoted or discouraged according to the technical and social affordances available within an online community. In most online communities, certain types of lying will be acceptable and others will not be. The owner of the community must consider how they can design the online environment to support the desired types of behaviour. In addition, they must decide how and whether to police deception. The following questions should be considered by a community developer when creating their strategy towards online deception.

1. Is the purpose of the strategy predominantly regulative or investigative?
2. What are the social characteristics of the community?
3. What affordances for deceptive action are present in the technology?
4. What types of lying are likely to be prevalent or particularly detrimental within this environment?
5. What investigative detection techniques are the most suitable for the community (as characterised within steps one to four)?

### 5.2.1 Regulation and Investigation

A distinction can be drawn between investigating a particular individual based upon reports by a third party, and simply monitoring a community for the purpose of identifying deceit. In the former case, it is likely that the investigator is already aware of the type of lying being employed. Contrastingly, for the purposes of regulation the investigator must focus their effort according to the type of lying that is most prevalent or damaging to the community.

### 5.2.2 Examination of the Social Context

The investigator must examine the social characteristics of the online community with a special emphasis on deceit. Some questions are central to this examination, including:

1. Are the majority of participants regulars or transients?
2. What is the motivation for interacting within the group?
3. How might a participant benefit from deceptive success?
4. What are the demographic characteristics of the group?

The answer to (1) might give an indication of the time and effort that the majority of participants have invested within the group. This would affect the interactive behaviours of deceivers who would need to operate either within or outside of established social networks. Understanding (2) and (3) would help to identify what some of those implications might be. The demographic characteristics of the group would affect the prevalent style of communication. Departures from

these patterns that seem 'out of character' might then provide a route towards regulative lie-detection.

### 5.2.3 Examination of the Technological Context

The investigator must define the nature of the particular CMC system under analysis. The vast array of features that can be implemented within CMC software reject anything but an inchoate classification. However, key characteristics that were identified in the previous chapter are:

1. Is it a synchronous, quasi-synchronous, or asynchronous system?
2. What level of anonymity is afforded by the system?
3. What level of interactivity is afforded by the medium for different types of user?
4. What expressive structural affordances are present for different types of user?

The synchronicity of the system (1) will indicate the extent to which a deceiver is pressured to manage conversational demands during the performance of their interpersonal deception strategy. Anonymity (2) is a key affordance for a deceiver who wishes to occlude their identity. The potential for asynchrony in the level of interactivity (3) afforded between different users can put deceivers at an advantage during their performances. Similarly, differences in the expressive functions can enhance the eloquence of some whilst putting others at a disadvantage. This is relevant because eloquence is a key rhetorical tool of the effective deceiver.

### 5.2.4 The Prevalence of Innocuous or Detrimental Deceit

When developing a lie detection strategy, it is important to focus on those types of lying that are of the most concern. Social lies, role-play, and identity exploration may be too innocuous to be of concern in many types of community. To challenge all instances of lying would be detrimental to the community, and so a technique should be designed such that it is easy to disambiguate everyday social lies from those identified as being particularly detrimental and hurtful.

### 5.2.5 Selecting a Detection Strategy

The questions posed in the previous steps provide a means of characterising the deceptive context. From these, a set of propositions may be derived, concerning potential deceptive behaviours that can guide the investigator in the selection or development of an appropriate detection strategy. Such a strategy may be unitary or composite (comprised of a selection of complementary techniques). A strategy can also be considered either active or passive. Passive strategies require the investigator to collect data, whereas active strategies require interaction with the target themselves. A strategy may involve holistic or specific techniques. Specific techniques can identify lying at a lower level of granularity than holistic techniques. Specific

techniques pinpoint lies whereas holistic techniques identify only the very existence of lying. A strategy can be overt or covert and can be invasive (such as through the use of surreptitious software) or non-invasive. Strategies can aim to identify individual liars (individual), or to gather a more high-level understanding of the prevalence of deceit within the online CMC context (social).

### 5.2.6 Present Difficulties

The process charted above is a framework for thinking, rather than a recipe for highly predictive lie-detection. The difficulty facing a practitioner is that an effective online lie-detection technique does not exist at present. However, the issue of prevalent deceit is current, and it is not helpful to delay constructive action pending the development of such a diagnostic tool. To this effect, investigators can still focus their efforts on limiting the affordances available to a deceiver over the other members of the community. Such decisions can be based upon an examination of the strategic behaviour of liars, and the ways they use technology to further their agendas. The role of strategic behaviours has been covered under the umbrella of Interpersonal Deception Theory. This theory also considers the role of the non-strategic leakage behaviours that may eventually provide a route towards the development of naturalistic CMC lie-detection techniques. It would appear that the application of IDT to CMC is a strong basis upon which to set off in search of these techniques. The rest of this chapter works towards this goal.

## 5.3 The Interpersonal Deceptive Process in CMC

The propositions of Interpersonal Deception Theory were presented in Chapter 2. One of the key postulates of this theory is *the principle of interactivity*. This principle holds that deception as an interpersonal process can be affected by the degree of context interactivity afforded by a particular medium. Burgoon and her colleagues (2003) have recently attempted to examine this relationship further, through the explicit investigation of deceptive CMC interaction. The authors suggest that the reduction of visual cues (and subsequently social cues; c.f. Filtered Cues Theory) would imply that CMC has a relatively low degree of context interactivity. If that were the case, then IDT would posit subsequent reductions in trust and an increase in skepticism during deceptive episodes (ibid). As well as being beneficial to a receiver, a sender might also profit from lower context interactivity. These benefits include a reduction of the cognitive load away from the regulation of nonverbal channels, heightened impression management (c.f. Hyperpersonality Theory), and less channel discrepancies (ibid). In addition to this, they suggest that reductions in trust may be overridden by an adherence to group norms when group identity is made salient (c.f. SIDE Theory). In this manner, Burgoon and her colleagues (2003) suggest that the role of CMC in a deceptive interchange is paradoxical. They designed a zero-history group study in order to investigate whether participants demonstrated less involvement, mutuality, and lower attributions of trust and truth under the CMC condition. These hypotheses were not supported. The authors note that although context interactivity is affected by the

absence of visual cues, this is not the only factor affecting interactivity, and more research needs to be performed in establishing what technological aspects of communication media contribute to interactivity as well.

Burgoon and her colleagues (2003) are careful to place their study in context, both in terms of the type of groups involved (zero-history) and the type of CMC system used. IDT and the investigation of context interactivity can be expanded further to more fully encapsulate naturalistic CMC deceptive episodes. The purpose of this section is to expand upon this discussion, which is centered upon the following points:

1. That the conceptualisation of CMC as an uninteractive medium in the filtered cues tradition contradicts more contemporary theories such as SIP. This has lead to a false theoretical division between the skills of a sender, the temporality of interaction and the interactivity of the context.
2. That a greater degree of asynchrony exists within deceptive encounters through CMC than is present face-to-face.
3. That the dyadic model of interpersonal encounters can be expanded into a group-based model.
4. That the reconstruction of speaker roles militate the superordinate effects of interpersonal orientation within a deceptive encounter.
5. That the role of nonstrategic leakage cues to deceptive behaviour within CMC has yet to be fully investigated.

These points are expanded upon in the sections that follow. The following summary of context interactivity within IDT has been provided as an aid to the reader.

<i><b>The Components of Context Interactivity</b></i>	<i><b>The Affects of Low Context Interactivity</b></i>	<i><b>Other influences that work alongside context interactivity</b></i>
<ul style="list-style-type: none"><li>◦ Access to social cues</li><li>◦ Immediacy</li><li>◦ Relational engagement</li><li>◦ Conversational demands</li><li>◦ Spontaneity</li></ul>	<ul style="list-style-type: none"><li>◦ Lower Initial Strategic Behaviour</li><li>◦ Lower Initial Expectations for Honesty</li><li>◦ Lower Initial and Ongoing Judgments of Sender Credibility</li><li>◦ Higher Initial and Ongoing detection accuracy</li><li>◦ Higher Initial detection apprehension</li></ul>	<ul style="list-style-type: none"><li>◦ The Goals of the Sender</li><li>◦ The Positivity of Relationship Between Sender and Receiver</li><li>◦ Familiarity Over Time</li><li>◦ The Skill of the Senders</li><li>◦ Deviations of Sender Communication from Expected Patterns</li><li>◦ Truth Biases</li><li>◦ Receiver Decoding Skills</li><li>◦ Expression of Disbelief</li><li>◦ Reciprocity</li></ul>

Table 5.1: Context Interactivity and IDT (adapted from Buller and Burgoon, 1996)

### 5.3.1 Context Interactivity and Social Cues

Those hitherto unexplored intricacies that complete the relationship between context interactivity and CMC do not result *solely* from the utilisation of technological artifacts that inherently act to either promote or encumber interaction. SIP theory has implicated that interactivity is a function of two other components within CMC; the skills of the collocutors and the duration of their relationship (Walther, 1992). Similarly, ecological approaches define the interactive potential of an artifact in terms of the affordances of that artifact *for a particular user or type of user* (Gibson, 1979; Hutchby, 2001). Whilst IDT acknowledges the roles that these two components play within interpersonal deceptive episodes (e.g. postulates eight and nine), it separates them apart from the definition of context interactivity. For example, for a person who cannot speak, the telephone affords a lower level of context interactivity than an email system. The telephone's inherent 'context interactivity' does not act to concurrently mediate that advantage; it is simply the case that the telephone does not afford a greater level of context interactivity over CMC *for a person who cannot talk*.

The counter argument to this claim, which has been leveled at many ecological approaches (see the discussion in Chapter 4), is that the relevant affordances of a particular medium are similar for a wide variety of people. If this is the case, then it will be easier to assume a stable level of context interactivity across a particular group or culture, than to worry about irrelevant individual differences. However, whilst there is some truth to this argument, the affordances available within CMC probably *do* differ within distinct social or demographic groups. People differ in their abilities to type, to operate computers, to express themselves verbally, to express themselves within the culture of the medium, and in the various software and hardware systems they use to interact within a CMC network. The access to social cues that is a key determinant of context interactivity is not simply lower in CMC, but is itself a function of the affordances of the medium for a particular user. Therefore, the social and technological skills of that user are partially expressed by the affordances that are available to them, and the ways in which those affordances are selected and managed. It is a peculiarity of CMC interaction that context interactivity might actually be different for different actors within a deceptive episode and that this disparity is more likely to be greater online than it is offline.

So the context interactivity of the medium is partially determined by the access to social cues which itself is determined by the mutual affordances available to a user. Skill is then redefined as having both a social and technological element, which is related to how those affordances are used. However, the access to social cues is not the only component of context interactivity that is relevant when the deceptive context moves towards CMC. Whilst the other components of context interactivity may vary within the medium in their own right, some are particularly tied to CMC interaction over face-to-face interaction. For example, conversational demands within CMC are relatively higher and transmogrified forms of turn taking can increase the pressure for spontaneous interaction. Therefore, whilst CMC cannot not be defined as a context that reduces the access to social cues 'across the board', some systems might still be characterised by a low

level of context interactivity because of widespread effects on immediacy, relational engagement, conversational demands, or spontaneity.

### **5.3.2 Perceptual Asynchrony**

In a paper examining the sender, receiver, and observer's perceptions of reality during deceit, Burgoon and colleagues (1996b) surmised that senders and receivers jointly construct the reality of a deceptive episode and have similar perspectives concerning the perceived veridicality of the exchange. In Chapter 4, it was revealed that within online communities reality could be constructed either by groups of participants or by the individuals themselves. Consider Turkle's (1995) categorisations of online relationships as being either virtual with the virtual, virtual with the real, and the real with the real. Virtual-with-the-real relationships are common within CMC and can result in hurtful consequences. Online interaction can be known to result in relationships that are perceived by each actor as having a different level of seriousness. For example, a deceiver may not treat a particular relationship seriously and this attitude would affect their expenditure in the strategic activities that are central to the propositions posited by IDT. According to proposition seven, of IDT, this disparity in the perceived serious of the interaction will cause an asynchrony in the strategic displays of the sender and receiver. This will in turn affect subsequent interaction. The key role of CMC is to provide opportunities for this sort of asynchronicity to occur more frequently than it would in face-to-face interaction.

Perceptual asynchrony may also occur because people use different types of software or computer setup to access an online community. Users can configure their chat software to remove features such as pretty-printing and may have different versions of the client programs. Such differences represent an asynchrony in the affordances available between senders and receivers, similarly constraining action and perception. Indeed a well-motivated liar may actually set out to attain the most recent or most functional version of the software to be at an advantage over the receivers.

### **5.3.3 Multiplicity of Interaction**

IDT adopts dyadic face-to-face communication as its exemplar. However, interaction within a group is perhaps the more widespread context of online deceit. CMC chat rooms can be very public places where conversation can involve a large number of different speakers. Even dyadic communication within these forums has an audience to which a sender may need to respond. The co-construction of a deceptive episode amongst a group of people has yet to be explored at the interpersonal level. The question remains how such encounters differ from the dyadic approach so that IDT might be socialised.

The most basic group scenario is where there is only one deceiver (sender) and a multitude of receivers. As with the dyadic model, receivers will have suspicions that are displayed and that result in behavioural adaptation by the sender. However, receivers in a group are more likely to co-construct and co-ordinate these displays. In accordance with SIDE, where group identity is

made salient, adherence to group norms will result in the co-construction of suspicion displays. Where individual identity is made salient, suspicion displays will be more individualised. In these two different scenarios, it is unclear what the results may be. Each of these will now be discussed in turn.

#### **5.3.3.1 Salient Group Identity**

Where group identity is made salient, adherence to group norms can promote a high level of consistency in the exhibition of suspicious behaviour. This agreement need not be made verbally, but is likely to be exhibited through behaviours that reflect high group cohesion. Such group displays of suspicion are likely to be driven by the more socially skilled members of the group. It is likely that deceivers will similarly attribute most of this suspicion from the more socially skilled people. This attribution would occur because of the expressive salience of these people (a similar phenomenon is found with physically salient targets and has been dubbed the illusory causation effect; McArthur, 1980). Subsequently, deceivers may increase their strategic behaviours towards these people if they believe that they can rhetorically compete with them. However it seems more likely that the experienced deceiver would try to divert attention away from the socially-skilled people if they believe that they cannot compete with them or simply because it is easier for them to do so.

The group to which the deceiver affiliates is also relevant. If the deceiver has previously been a part of the group, then suspicions may be mediated by the prevailing perception of in-group trust. Conversely, if the deceiver is a member of an out-group, suspicions may be militated by lower levels of trust. Campbell (1967) proposed that such intergroup stereotypes arise because they reflect a group member's actual experience. As out-group relations are external relations, they are naturally confronted with greater distrust (Brewer and Brown, 1998).

#### **5.3.3.2 Salient Individual Identity**

Where individual identity is made salient, people are more likely to conform to personal attitudes and perceptions. Therefore, there is likely to be an asymmetry in the degree of suspicion displays exhibited by different receivers under this condition. The deceiver may choose to interact more heavily with those receivers who display less suspicion so long as this does not conflict with the goals of the deceiver. As suspicion displays are related to the social skills exhibited within the medium, less skilled receivers (and less socially skilled people) are more vulnerable to deceit because they are less competent with the medium, and are less likely to exhibit suspicion.

### **5.3.4 The Reconstruction of Speaker and Listener Roles**

The peculiarities of QS-CMC mean that a person cannot construct an utterance whilst also observing the interaction around them (Garcia and Jacobs, 1999). During utterance

construction, it is possible that a collocutor will send a new posting that would alter the way that the partially built utterance would be interpreted. Similarly, the new posting may make the partial utterance redundant, requiring a reformulation of a response from scratch. If a sender decides that they wish to strategically alter their utterance, they are presented with a dilemma because editing takes time, and the speed of interaction is an important determiner of power. For a deceiver, the dilemma exists between two different types of strategic power, that of speed and that of content. In a forum where speed of response is contested, the deceiver may be at more of a disadvantage because of the less time spent reviewing utterances.

### **5.3.5 The Lack of Research into Non-strategic Leakage**

Non-strategic leakage within face-to-face contexts has been well researched under the four-factor paradigm (Zuckerman et al., 1981). IDT suggests that strategic behaviour often supercedes that of non-strategic behaviour, reducing its relevance in everyday deceptive encounters. However, the non-strategic leakage cues within the CMC context have yet to be researched to anywhere near the same extent. The suggestion that CMC allows fewer non-strategic cues could be premature, and to generalise that strategic cues within CMC have the same power as they do within the face-to-face context could also be premature. These cues are discussed further in the following section.

## **5.4 The Non-strategic Deceptive Behaviours of CMC**

The interpersonal deceptive process outlined in the previous section refers to both strategic and non-strategic displays. Non-strategic 'leakage' displays have been studied at length throughout deception research, and are driven by the four factors that were presented by Zuckerman and colleagues (1981). The non-strategic cues to deceit are usually considered as nonverbal, paralinguistic and physiological expressions of psychological states that correlate with deceit. The most consistent findings are found with indicators of concealed emotional expressions and autonomic arousal. Some of these findings include:

- *Facial asynchrony as autonomic muscular control of the face under certain emotions conflicts with the strategic control (Ekman, 1984)*
- *Increased blinking when trying to suppress negative emotional affect (Gross and Levenson, 1993)*
- *Exhibition of very short 'microexpressions' that reflect the true emotional affect before they are suppressed (Hagaard and Isaacs, 1966; Ekman, 1984; Ekman and Friesen, 1984)*
- *Latency effects of recognition under the orienting reflex (Farwell and Donchin, 1991; Seymour et al., 2001)*

Within CMC, most of these cues are not promulgated through the medium. Nonverbal behaviours within CMC would seem to be constructed through purposeful action. The user



decides which font to use, the colours in which to display their text, the time they take to post, and degree to which they use pictographs. However, this does not mean that there are no non-strategic leakage clues within CMC because of two reasons. Firstly, verbal behaviours can have a non-strategic element (e.g. type-token ratios; see Chapter 4) and these automatic verbal exhibitions could hold the key to CMC lie-detection. Similarly, verbal lie-detection systems such as CBCA and RM propose that verbal behaviour when telling a story is dependant on the contents of episodic memory and as such may reflect the veridicality of the source. Secondly, nonverbal and physiological behaviours may affect a users control with a computer interface. It is possible that such changes are significant enough so that they can be detected by software.

When two participants are interacting through a mediated technology, a distinction can be drawn between those behaviours that never promulgate to the computer, those that only promulgate to the computer of the transmitter, and those that transitively cross the ether and are available to the collocutor (assuming that the CMC interfaces are essentially identical). Promulgated behaviours are potential candidates for the transmission of leakage displays of deception and suspicion. Those that do not pass the ether are still available for detection by the software system itself, but do not participate in the mutualistic relational processes predicted by IDT. Some of the cues that can be recorded by a CMC system can be broken down into the following broad categories:

#### **I. Not promulgated**

1. Micro-level interface behaviour (e.g. keypress and mouse click times)
2. Macro-level interface behaviour (e.g. scroll bar and window movements)
3. Utterance construction behaviour (e.g. utterance initiation time, deletion behaviours)

#### **II. Promulgated**

4. Utilisation of ideographs
5. Utilisation of pretty-print
6. Utilisation of other textual paralanguage
7. Verbal clues
8. Time between postings

These behaviours are exhibited both strategically and non-strategically. The purpose of this section is to speculate whether non-strategic leakage is passed on via these cues. This discussion will be structured around the four factors presented by Zuckerman and colleagues (1981).

### **5.4.1 Cognitive Load**

The cognitive perspective holds that because lying is more difficult than truth telling, valuable cognitive resources are diverted towards this activity and away from other areas. A deceiver

may need more time to plan what they are going to say, and this may lengthen the time it takes for them to start typing. Similarly, once an utterance has been constructed, more time might be spent in reading the utterance through, resulting in a longer time-to-post. It is also fair to speculate that the difficulty of constructing a lie may result in more revisions to the utterance and this will increase the editing time. For a deceiver, these lag times are indicative of processes that are superordinate to the processes of (1) maintaining good spelling, (2) using pretty-print, (3) using textual paralanguage, and (4) having a high level of verbal expressiveness. When a deceiver is in a context that is characterised by a pressure to respond, cognitive economising will result in less concentration on the aforementioned factors in order to allow more time to be spent on strategic action (IDT postulate five).

### **5.4.2 Affect: Duping Delight**

Some people may feel excitement when deceiving and this has been referred to as 'duping delight'. Such excitement comes from the rewards of deceit per se or from the rewards that are attained by using deception. The level of duping delight can vary between people and situations and Ekman (2001) provides three factors that may enhance this form of excitement.

1. If the lie is particularly difficult or challenging
1. If the receiver has a reputation of being difficult to dupe
2. If there is an audience who are aware of the deceit

These factors operate online in a similar fashion to how they might do offline (for example, a MUD wizard may have a reputation as being difficult to dupe). However there are some relevant differences. A deceiver within CMC may be in the physical presence of others who are aware of the deceit and this might enhance the duping delight. However, in this scenario, there is no need for either the deceiver or the audience to control their auditory or bodily behaviour, and they are free to physically and verbally express delight and excitement. Normally, the feeling of suppressed excitement associated with duping delight may result in signs of stress (Vrij, 2001) including (most pertinently), an increase in body movements. However, as the deceiver is not under the same degree of physical suppression within CMC as they are face-to-face, it remains to be seen whether these stress signals are still present. It might be the case that deceivers are more likely to become engrossed in the interaction when they *do not* have a co-present audience and therefore are less likely to express themselves physically. So perhaps in CMC, the lack of an audience might make duping delight easier to detect.

The cues of excitement that promulgate through the medium are closely linked to the cues of arousal discussed in section 5.4.4. In addition to these cues, excitement in general may act to speed up interaction, as the deceiver, like an addict, increasingly strives to attain the affective rewards of duping. This might then result in less sophistication and increasingly sloppy behaviour as the deceiver tries bigger and bigger lies in order to attain the same level of affect. Ekman (2001) tells a story of a man who stole books from a University bookshop. He increasingly stole larger and more difficult to conceal books until eventually his deceit became

so blatant that he was caught. Similarly, in the account provided by Quayle and Taylor (2001) in Chapter 4, the paedophile Q.X's behaviour became increasingly sloppier as he tried to attain the same level of arousal through ontological deceit.

### 5.4.3 Affect: Detection Apprehension

As with duping delight, under certain circumstances a deceiver will feel the fear of being caught (detection apprehension). Ekman (2001) states that detection apprehension will be highest where the reward for success is high, but the punishment for failure is also high. Within many CMC scenarios, it might be thought that there is very little scope for punishment because interaction is predominantly anonymous. However, it was shown in Chapter 4 that people could value their online self just as highly as their offline self and so punishment of that online self may be feared. Detection apprehension may also be high if the deceiver is engaging in illegal practices, such as paedophile-related activities.

Fear can similarly result in exhibitions of stress and as such can result in arousal cues. Fear may also cause a deceiver to be more careful about what they say, resulting in inhibition effects similar to those that result from an increase in cognitive load.

### 5.4.4 Arousal

The general physiological arousal of a deceiver can be hypothesised to occur because of one of three reasons suggested by Davis (1961). Conditioned response theory suggests that arousal would result from the conditioning of a negative emotional response with the content of the lie. Conflict theory suggests that arousal results from the internal conflict between the tendency to tell the truth and the need to lie. Punishment theory attributes the arousal to the fear of punishment resulting from the uncovering of the lie. As well as these three approaches, there is also an arousal response resulting from guilty-knowledge. The primary question for the arousal perspective of CMC lie-detection is to what extent does heightened physiological arousal result in skeleto-muscular effects that affect typing characteristics and other levels of control with a computer interface. These effects may be detectable by themselves, or may result in further effects that are reflective of the deceiver's adaptation of, and compensation for these influences. However, most arousal clues are unlikely to promulgate through the ether.

### 5.4.5 Attempted Control

Deceiver's will generally attempt some form of strategic action in order to appear more credible. These attempts can concurrently result in the leakage of non-strategic information (the specific strategies that might be employed by deceivers are discussed in the next section). One typical response is to adjust behaviour according to the *perceived* behaviours of veracity or in opposition to the *perceived* behaviours of lying (Zuckerman et al., 1981). This control can result in behaviour that departs from natural veridical behaviour because it is based upon a personal 'theory' of lying that is constructed on misinformation. The stereotypically perceived behaviours

of online deception were touched upon in Chapter 4, and are investigated further later in this chapter. Examples from the literature indicate that people are suspicious of 'a/s/l' requests, demands for cybersex, and controversial statements. Whilst attempted control can only indicate that the sender is attempting to look truthful (they may not necessarily be lying), this is one route by which deceit could be identified verbally.

Another potential leakage clue that falls under the domain of attempted control might include the deletion behaviour of the deceiver. During the construction of their utterances, QS-CMC users can engage in editing that is occluded from the collocutor. Deceivers may reveal contradictory information in these deletions that are easily recorded by software. Alternatively, due to the highly strategic nature of deceivers they may delete text more frequently, delete a larger chunk of the text, and rewrite utterances with a greater number of modifications.

#### **5.4.6 A Caution**

The point should be emphasised, that as with any nonverbal route to lie-detection, these non-strategic clues are not in themselves reliable indicators of lying. For example, behaviours associated with stress, cognitive load, or excitement can all result independent of whether or not a speaker is lying (Ekman, 2001). Therefore, the non-strategic route to lie-detection must involve a very careful observation of the deceptive context, and is likely to succeed only when non-strategic behaviours are observed that contradict the content of verbal behaviours and whereby the only explanation for this results from the presence of deceit.

### **5.5 The Strategic Deceptive Behaviours of CMC**

The strategic behaviours employed by chat room deceivers have been the subject of a very limited amount of research. Chapter 4 examined some instances of online deceit and these can be split into the following examples (1) lying about identity, (2) highly selective and inventive impression management in online relationships, (3) trolling, and (4) faking expertise. This study aims to explore deceptive behaviour in more depth, to confirm the presence of these phenomena and to uncover others.

The ideal method for such a research question is observation, however the observation of naturalistic deceptive behaviours is difficult because it requires knowledge of the ground truth. This can only be ascertained through intervention whereby the actions of the investigator could confound the phenomena under observation. Whilst ethnography into the deceptive practices within an online community would be a fruitful research topic, ethical considerations would probably preclude its implementation. Therefore, this study will seek to uncover the self-reports of chat room users who have experienced deceit.

Self-reports concerning deceptive phenomena are best evaluated within a constructivist framework, despite the fact that lie-detection research typically lies in the realm of positivism. Deceptive realities are co-constructed by both the deceiver and the deceived alike and whilst

these accounts are in many ways similar, qualitative differences do remain (Burgoon et al., 1999). This is especially the case within CMC, due to the perceptual asynchronicity issues discussed earlier. Therefore, it is necessary to attain accounts from both deceivers and the deceived alike in order to attain a more balanced view of deceptive exchanges.

In order to explore strategic behaviours in more depth, this study aims to answer the following research questions:

- How do respondents who have been affected by chat room deceit, describe the way in which this event occurred, the ways in which they became suspicious, the actions they took, and the effect that this deceit had on them?
- How do respondents who themselves have engaged in chat room deceit, describe the way in which the event occurred, the ways they tried to make themselves more convincing, the reasons why they might have detected suspicion, and the adaptations they made in response to this suspicion?

### **5.5.1 Method**

Seventy respondents, all of whom had been exposed to chat room deceit, took an online questionnaire that explored these experiences in greater depth. The respondents were recruited through CMC-relevant newsgroups, bulletin boards, and mailing lists. The questionnaire was constructed as a website (see Appendix I for more details) and contained predominantly open-ended questions. The respondents completed the questionnaire anonymously.

This design was chosen for the following reasons. As there has not yet been much published research on the activities of online deceivers, the purpose of this study was for qualitative exploration. To this end, the questionnaire was aimed at a very wide target population, in order to attain a broad selection of attitudes and perceptions from across the globe. In recent years, the World-Wide Web has become an increasingly popular tool for researchers interested in reaching a wide number of people without incurring the costs that come with conventional paper-based methods.

A questionnaire-based design was chosen over other qualitative techniques such as interviewing or ethnography. Lies and deceit are issues that are morally weighted in society and so it was expected that it would be quite difficult to attain responses from deceivers using an identifiable setting. Questionnaires can easily be made anonymous and the online environment can enhance this effect. However, questionnaires are not the ideal vehicle for attaining depth, as the questions cannot be modified to suite an individual participant and follow-up questions cannot be asked. In addition, the need to type a lengthy response into the computer, for very little reward produces inertia against the attainment of depth. This was deemed an acceptable shortcoming for an initial explorative study that is focused primarily on the breadth of deceptive experiences.

The possibility exists that malicious respondents would themselves report concocted fabrications. The ironic attraction of this form of ruse is difficult to miss, especially within an anonymous setting whereby the respondent does not need to confront the investigator. However, it was believed that because the questionnaire required a considerable investment of time to complete with very little reward, then this high cost-to-benefit ratio would deter this form of malicious attack.

The questionnaire was designed to be simple and easy to navigate. Due to the effort that a respondent would need to go to in order to complete the questionnaire, certain features were added to try to encourage responses. The participant was lead politely into the questionnaire with a clear set of short instructions that contained links to the questionnaire itself. Upon completion of the questionnaire, the participant would be taken to a page of interesting links to virtual communities and CMC-relevant sites as a simple 'thank-you'.

The questionnaire itself was split into four sections, and all of the sections were optional. The first section pertained to demographic data. The second section asked whether the participant had believed that they themselves had been the target of chat room deceit using the following questions:

- Q4. What was the other person trying to deceive you about?
- Q5. How did they try to accomplish this?
- Q6. Was there anything they said that made you suspicious, and if so what was it?
- Q7. Was there anything else that made you suspicious?
- Q8. What sort of impact would you say this deceit had on you?
- Q9. Please describe the kind of chat room that you were in at the time.
- Q10. Is this a chat room that you visit often, and if so how often?
- Q11. Are there any other details about this experience that you would like to add? If so please use the space below.

In the third section, the respondent was asked whether they themselves had engaged in deceit, and the following questions were asked.

- Q12. What were you trying to be deceptive about?
- Q13. Did you use any techniques to try and make yourself sound more convincing? If so, what were they?
- Q14. Was anyone suspicious of you?
- Q14a. If somebody was suspicious, how did you know that they were suspicious?
- Q14b. If somebody was suspicious, how did you react?
- Q15. Did anyone "catch you out"?
- Q15a. If they did, what do you think gave you away?
- Q16. Are there any other comments about this experience that you would like to add? If so please use the space below.

The perceived cues to deceit were examined in the fourth section but this data was omitted from the analysis<sup>1</sup>.

## 5.5.2 Analysis

A substantial amount of recruitment was needed in order to obtain a sufficient number of responses. The public, anonymous features of this design also made it susceptible to abuse, through the posting of malicious or offensive material. In total, seventy valid responses were received containing varying accounts of online deception. These accounts were grouped into categories of (1) Gender deception, (2) Age deception, (3) Relationship status, (4) Faked affection, (5) Invented vocation and skills, and (6) Faked origin or location. These themes will be discussed in the sections that follow.

### 5.5.2.1 Gender

The performance of gender deceit within chat rooms can be displayed by both men and women over a variety of time-periods, and this was reflected throughout the responses. The few accounts provided by 'deceivers' described episodes of a short, 'harmless' nature carried out for light entertainment, as one male respondent writes *"I made a friend look a fool by convincing him I was a girl"*. Another male respondent reports having 'gone undercover' in order to trick a friend (located in the same room) who was attempting to meet girls online. The respondent assumed the identity of one such 'girl' and ironically it turned out that their friend was also lying about their appearance. Certain stories reported by victims of gender deceit were not as innocuous and one involved an attempt at suicide.

One respondent (Martina<sup>2</sup>) tells a story that occurred in a close-knit, online community of adults in their twenties and thirties. Within this community, people often met in real life, exchanged phone numbers, and knew extensive information about each others offline lives. At the time of the incident, Martina was using the chat room for between four and ten hours a day, five or six days per week. One of these community members was named Chris (alias 'Thunder'). Martina and Chris became quite close as friends and exchanged intimate information between one another including 'provocative photographs'. It was normal for people within the community to meet offline but Chris always seemed to have an excuse that prevented him from meeting Martina, despite living relatively close. Eventually, Martina became suspicious, especially when watching how Chris was acting within the chat room.

*"He would often chat up the other young ladies in the room, flirting with them openly and in general, acted like a "guy" [...] My suspicions were piqued when every so often, during open cybersex in a room "he" was in with a young woman,*

---

<sup>1</sup> This section contained quantitative scales that assessed the degree to which respondents believed certain behaviours to be indicative of deceit or veracity. The unexpectedly small number of respondents precluded the statistical use of this data.



*Thunder kept referring to his genitals as his 'package'. He referred to female genitalia in very gentle terms...there was just something about the technique that made me think he was a bit feminine."*

Martina had Chris' home address (they had exchanged letters) and decided to pay him a surprise visit. As Martina writes, *"Thunder turned out to be a 250 pound lesbian"*. The outcome of the deceit was to fragment the community, many of whom no longer speak to one another. Martina speaks further about the effect on one of Thunder's online 'lovers':

*"She literally left her husband and kids, to move to be with 'him'...to have a future with the person she believed was the MAN she loved. She tried to commit suicide shortly after the deception was revealed to her... Thankfully, her attempt fell short of success."*

As with the story of Alex, reported in Chapter 4 by Van Gelder (1983), Thunder attempted to maintain relationships after the deceit had been unveiled, justifying this by downplaying the importance of physicality. For Martina, however, this was unacceptable and the moral weighting she ascribes to this action is clear from her response:

*"She failed to see the reasoning that the foundation of our friendship was built on a lie. That her deceptions ran so deep, that there was NOTHING left to salvage. She refused to see beyond that...I felt I could not continue to keep this person in my life. In a strange way, it felt like a death. My "brother" Chris was gone, only to be replaced with this heinous human being who could not understand the concept of truth."*

Other victims of gender deception also reported being suspicious of language-use that was either stereotypical or atypical for the gender portrayed by the sender. One respondent reported that the 'thirty-year-old woman' they were speaking to seemed to be obsessed by sex but demonstrated a lack of any real practical understanding of sex itself. Stereotypical behaviour was often referred to as 'trying too hard'. There was some reporting of other simple heuristics being used to detect deceivers. Refusals to send pictures, to provide real world information, or through 'making excuses' as to why a relationship cannot be taken offline also induced suspicion.

In one chat room focusing on transsexual issues, a participant reported the prevailing distrust of newcomers using the A/S/L opening sequence. Such requests were seen as being indicative of what the respondent calls a snert, *"someone who is in an inappropriate chat room or IM with intent to harrass, receive sexual gratification, threaten, and otherwise cause harm through deception"*. As is the case with trolls, so called 'snerts' are generally ignored within the group.

Little information was provided by 'deceivers' as to their tactics or behaviours but some of these were reported by the targets. As with the accounts of paedophilic deception reviewed in Chapter

---

<sup>2</sup> Names and pseudonyms of people referred to within this analysis have been changed.



4 (and the story of Thunder), deceivers often used photographs of friends, family, or celebrities in order to add credence to their masquerade. One respondent reported being deceived by a man who used a picture of his girlfriend to support the false information provided in his profile.

A unique quality of CMC is that a single faked identity can be managed in real time by more than one person. A situation was reported of two men who jointly participated in the creation of a single female identity whilst at an Internet cafe. It can be speculated that this form of cooperative deceit can help alleviate the cognitive difficulties in managing a fabricated personae leading to a more effective performance. Another person reported easing these difficulties through the adoption of a close friend's identity. One man reported 'successfully' participating in a lesbian chat room under the assumed identity of a woman, constructed by amalgamating the traits of two female friends.

#### **5.5.2.2 Age**

Age is another characteristic that is performed verbally. One respondent reported having talked to a 'teenage mother' who was apparently only fifteen-years-old and had two children. This in itself raised suspicion with the respondent, despite such a situation being perfectly plausible. This raises an interesting point relating to the diversity of people met online. Receivers are exposed to a wide variety of different people, from different cultures and social groups, especially considering the attraction of CMC for demarginalisation. However, receivers may still maintain locally derived perceptions concerning the cues to deceit and these may include behaviour that they consider atypical. This may be one of the main contributing factors to distrust within the medium. This was not the only factor used by the respondent however. Whenever the girl sent pictures of herself, no children were present, and when one of her offspring was reported to have died *"she didn't seem too upset"*.

Another respondent suspected that a person she was speaking to within a chat group for people in their twenties was not as young as he appeared to be. The respondent believed the man was attempting to appear fifteen years younger than he was. The respondent became suspicious when the man did not have very much to say from his time at school, something that a twenty year old would have a better memory of and a more contemporary story to report. As is quite common throughout the transcripts, the man sent a 'picture' of himself to the respondent, and as she writes, *"The picture was very nice and he looked like he should have been a model, not talking to a stranger at 2 am"*.

These two examples concerned people trying to appear younger. Where people were attempting to act older, as with instances of gender deception, receivers were made suspicious by 'overcompensation' and 'trying to hard'. One instance of this puzzled a respondent who said that they were offended that the other person did not reveal their age, as they did not really care what age they were. Other respondents became suspicious when apparently mature adults revert to teenage slang or 'immature' language.

One respondent mentioned being lied to by a thirty-year-old journalist who was claiming to be only nineteen. The journalist apparently was fairly convincing because of the way she interlaced facts with lies. The revelation followed a refusal by the respondent to engage in cybersex, whereby the journalist confessed the following day. Although they still maintained their friendships for a few weeks, the respondent felt disappointed and eventually they lost contact.

### 5.5.2.3 Relationship Status and Affection

Some of the targets in the study related stories of having built up close relationships in chat rooms only to find out that their online intimates were already in relationships in the offline world. Sometimes these involved simple lies of omission, providing hints using the incorrect inference dodge (Ekman, 2001). One participant recalls being led to believe that the person they were talking to had broken up with her boyfriend, *"She did not 'say' this, but rather gave vague hints and seemed to want me to make my own assumption"*. The respondent became suspicious when he noticed that 'her story' was inconsistent, and various different versions of events were being told. Additionally, he reported a suspicion of the manner in which the woman asked him to address emails, *"I had to use her name in the subject line of any emails I sent her - it was obvious that she had set up outlook to dump those emails in her account so her b'f did not see them"*.<sup>3</sup> Whilst such behaviour might be completely innocent way of sharing or managing an email account, it would seem as though this request provoked suspicion because of the affordances that it presented the woman.

Another respondent reported an instance of betrayal by their online male partner with another woman online:

*"he said he loved only me (online of course) [...] by very detailed chatting, romantic emails, constant tries to make me think he only chatted romantically with me, not others"*

The façade was broken when the respondent saw the way in which the man acted with another woman in the chat room. The respondent also thought that the man seemed to try 'too hard' to please her, obviously with some success:

*"i know how silly it may sound but you become VERY CLOSE to someone when chatting with them over long periods of time, day after day. true intimacy does develop and lying when you have a relationship of that depth hurts, just as it does outside of chat"*.

Many other victims of similar episodes reported that these sort of incidents was destructive on their willingness to trust people through the medium, making them somewhat disappointed and jaded, as one respondent writes *"it is bad when there is no where a person can go to where someone could be honest, we stay on the net cause we have [...] been wore out by all the lies from people in the real world"*.

#### 5.5.2.4 Vocation and Skills

Lying about having an interesting job or a set of skills may be appealing due to the positive feedback received by the listener. Such a process may be critical to the construction of a 'perfect identity' whereby a person may initially be freed from their 'shortcomings' and briefly experience the sort of admiration that they perceive others to have.

In MUDs, this sort of identity reconstruction is common, and indeed part of the reason for participating in the first place. One MUD implementer recalls meeting a person who he describes as a pathological liar. This person seemed to be telling somewhat far-fetched stories about their abilities in the martial arts, having apparently spending years training with masters in Japan. This person related a story surrounding having 'killed someone' whilst training, as the respondent reports:

*"He attempted to convince me that he killed somebody using martial arts mostly by expressing remorse that it had happened, [...] I was not convinced and said so, and he mocked me in a somewhat smug, patronizing way, saying that he didn't expect me to fully understand what he was talking about because of my "western" way of thinking - saying, in effect, that my disbelief was possibly related to my not-as-expanded mental faculties. [...] There was not necessarily one phrase or pattern of words [...] - it was mostly the story itself, the un-believable-ness of the content of this story."*

What makes this interesting is the presence to handle an accusation by attacking a listener's fundamental belief set, instead of by addressing the specific incident, performed to diminish the credibility of the listener and make them doubt their own perceptions. Another respondent conjectures as to the reasons people lie about their abilities online:

*"people seem to want to parade thier lives in front of you I have found that they are usually 'offering up to market' and by that I mean tempting you with stories that if you do not give them 'credit' for then they will raise the stakes and try to impress you more"*

MUD administrators sometimes play down their abilities and assume a false identity in order to covertly roam their own MUD. One administrator described some techniques that he used:

*"pretended to be \_very\_ gullible, believing things which I - in reality - saw through in an instant. I also made deliberate typos and grammar mistakes, and even used different emoticons than I usually do."*

This latter remark, concerning emoticons implies that this is something to which the administrator believed that other users might be sensitive. The administrator chose to use the smiley :) as opposed to the more contemporary alternative :-).

---

<sup>3</sup> Microsoft Outlook™ is a popular emailing client.

One respondent reports having been lied to by a man who was apparently a Chemical Engineer. The respondent became suspicious upon talking to friends who really were chemical engineers and who discredited some of the claims provided by the man. Sometimes such vocational deceptions are far more blatant, one respondent reports a man telling him he was an 'educator', but the man's poor spelling evoked suspicion.

#### **5.5.2.5 Appearance**

Respondents that had lied about their appearance tended to have more defensive rather than aggressive motives. One respondent who describes herself as 'large and awkward' had initial reservations about disclosing her appearance and lied about this when she started to talk to people online. When she was asked for a picture she would send one of a family member (who actually looked quite similar to herself) as she explains:

*"It was a ridiculous notion...to this date, I do not understand why I did that. I suppose the notion of being able to create a "new" me sound appealing"*

She was surprised to find that when she did eventually post a real photo of herself the other people in then chat room were actually quite complimentary, more so than she would have thought in the beginning.

Another respondent mentions a tactic they used that played on the perception of the prevalence of online deceit, *"when first asked what I looked like, I said I was fat and ugly, then said I was joking and gave them the description I wanted them to believe"*.

### **5.5.3 Evaluation**

As with face-to-face encounters, there exists a diversity of deceptive phenomena within CMC and a varied subjectivity concerning the morality of that behaviour. The playful deceiving of complete strangers seems to be an acceptable and justifiable activity for some, heralded under the slogan of freedom from the body and emphasising that chat room users should be responsible for their own prudence. The responses revealed that the deceiver and the deceived often have very different impressions as to the moral weighting of the deceit. As has been mentioned throughout this chapter, perceptual and other forms of asymmetry (c.f. Turkle, 1983) seems to be an influencing factor within deceptive episodes.

It also seems as though sender deviations from 'typical' behaviour is a key determinant underlying suspicion, whilst stereotypic behaviour is also mistrusted. This can erode the perception of trust within CMC in general, where interaction can occur with a wide variety of different people adopting a diverse range of behavioural norms and stereotypes. This is militated by the finding that under some circumstances, behaviour within CMC has a tendency to be highly stereotypic (c.f. SIDE).

Some techniques have been uncovered that deceivers use in order to appear more convincing. These include the joint production of online personae, playing upon stereotypes, sending photographs of other people, and telling receivers 'what they want to hear'.

This study has explored a few of the ways in which deceptive episodes have been carried out online and is one of only a handful of studies to attempt this explicitly. This study is limited by its exploratory nature, low sample size and breadth of focus<sup>4</sup>. However, important themes have been uncovered and one area for future research stands out. It would seem as though the role of stereotyping is a key factor in strategic deceptive encounters online and the examination of this interplay is likely to be a critical component driving a potential naturalistic lie-detection technique.

## 5.6 A Semantic Scheme for QS-CMC Lie-detection

In Chapter 2, it was suggested that any categorisation scheme for lying and deceptive behaviour would primarily reflect the priorities of the classifier. The investigator interested in CMC-based lie-detection is best served by adopting a scheme that is appropriate to that particular endeavour. This system will reflect the types of deceit that are common within CMC, but also the ways in which lie-detection studies are conducted. A proposal for such a semantic scheme is presented here.

It is useful to think of an *interpersonal deceptive strategy* as encapsulating a deceiver's strategic agenda. The *lying behaviour* within that strategy is comprised of a sequence of interlaced utterances that are either *perceived truths* or *lies*, both of which need to be scrutinised for the purposes of lie-detection. Lying behaviour can be classified according to five pertinent dimensions of *content*, *perceived stake*, *temporality*, *elidity*, and *relationship*. The stake of the behaviour is a binary classification representing whether the deceiver has much to lose resulting from the revelation of their deceit. As has been shown in previous chapters, stake is usually indicative of the affective pressures to which that the deceiver is exposed. The temporality of the behaviour can be classified as either single event, multiple event, or as an ongoing exchange. The research presented in Chapter 5, showed the significance of temporal factors and how they effect interaction within CMC. The elidity of lying behaviour refers to whether the strategies employed are predominantly commissive or omissive. Finally, the relationship between the deceiver and the deceived is also an important factor in assessing the how the deceptive episode is constructed by the collocutors.

Lies can be split into four content categories (Ontological, Epistemic, Episodic, Affective) that reflect the predominant type of information that is being lied about. Ontological lies are those that mask identity or being. They are exemplified by the sort of identity reconstruction reviewed in the previous chapter. People can lie about their physicality, personality, socio-cultural

---

<sup>4</sup> A more comprehensive evaluation of this and other research described within this thesis has been provided in Chapter 9.

characteristics or any other aspects of their being. Conventional techniques, such as polygraph tests, have been used to detect ontological lies, but no unique theoretical distinction has yet been made that identifies this form of lying as being grounded in its own unique psychological process. This distinction would be useful because of the ambient presence of ontological lying within online communities. Epistemic lies are those that claim knowledge where no such claim is justified, or that provide false knowledge that does not fall under any of the other content categories. Episodic lies are those that involve the fabrication of events (episodes), and are the focus of techniques such as Reality Monitoring. A person may use episodic lies to provide a history that is consistent with the ontological aspects of their deceptive strategy. Affective lies conventionally examined nonverbally or physiologically, are those lies that mask or masquerade emotion. As was shown earlier, this form of lying represents a particular challenge to the CMC domain where conventional nonverbal cues are elided.

## 5.7 Summary and Discussion

This chapter began with a list of five objectives. Firstly, a practical process for considering lie-detection within an online community was outlined, and it was explained that lie-detection in this form was embryonic and so the community developed must restrict their activities to limiting those socio-technical opportunities that their CMC systems affords deceivers. Then, in order to address some of these shortcomings, the application of IDT to the domain was revisited and key issues were highlighted that expanded upon the adaptation recently suggested by Burgoon and colleagues (2003). Further to this discussion, some non-strategic cues to CMC deception were put forward by considering the relationship between user-interface control and behaviours that correlate with deceit. Research considering the strategic behaviours of deceivers was performed in order to provide more information about this unexplored area of CMC research. It was proposed from these findings, that stereotypical behaviour in CMC (partially driven by the de-individuation effects posited by SIDE) could be a key factor affecting suspicion and deceptive displays. Finally, a semantic scheme was designed, inspired by the existing lie-detection literature and the research carried out on strategic cues, in order to be a system suitable for discussions of CMC lie-detection.

This chapter has highlighted a number of issues concerning naturalistic lie-detection within CMC, and as was proposed earlier, much of this research is relevant to the field of controlled lie-detection. This thesis addresses controlled lie-detection in the next three chapters. In Chapter 6, CMC deceit is investigated from two perspectives. One non-strategic route towards controlled lie-detection is proposed and tested by examining whether a particular psychophysiological effect (the orienting reflex) can be detected through response-time changes at a conventional computer keyboard. Secondly, it is investigated whether or not knowledge of a CMC-based crime is salient enough to be detected by such a technique. In Chapter 7, it is assessed whether or not an investigative interview can be carried out entirely through CMC and the advantages of doing so are proposed. In Chapter 8, this interview is



modified and used as a part of a larger CMC-based Statement Validity Assessment, testing the applicability of content criteria and investigating some of the non-strategic leakage clues proposed in this section.

## Chapter 6

# Response Time Measures to Assess 'Guilty Knowledge' within a CMC Scenario.

The effect of guilty-knowledge on mouse-button response-times is assessed using a variation of the 'oddball scenario' described by Farwell and Donchin (1991). Participants were asked to conduct a novel form of mock-crime within a computer-mediated chat room. Within the chat-room they were asked to cause a disruption to a group task that was being undertaken by five other volunteers. The chat room was only a simulation but post-hoc interviews revealed that this manipulation was convincing. The participants were asked to complete two lie-detection tests, a guilty test and an innocent test that assessed their guilty-knowledge of information that was relevant to the mock-crime. Each test required the participant to click the left mouse-button if they recognised a 'target' word-pair (learnt specifically for the test), and to click the right mouse-button for any other word-pair. The participants were told that they must click the appropriate mouse button within one-second of being presented with the stimulus. In amongst the irrelevant word-pairs within the guilty test were mixed six items that related to the mock-crime. During the guilty test, the participants were approximately 100ms slower in their responses to these crime-relevant items than to the other irrelevant items. This difference was not found for the innocent test. The results were similar to those reported by Farwell and Donchin (1991) but considerably more conservative than those found by Seymour et al. (2001). The results provide evidence that low-level reflex actions that correspond to guilty-knowledge can be detected using standard, everyday PC apparatus. This could mean that physiological leakage clues might propagate themselves through the medium and this could provide channel for CMC lie-detection. Additionally, the investigation demonstrates that similar CMC-based 'crimes' are salient enough to be assessed using conventional lie-detection techniques such as GKTs.

---

### 6.1 Background

Little is known about the relationship between deceptive behaviour and human-computer interaction. However, research has revealed that there are a number of behaviours indirectly associated with deceit that are the product of motivational, affective, autonomic, or cognitive factors (Zuckerman et al., 1981). These behaviours can be expressed either nonverbally or physiologically. A central question for the field of CMC lie-detection is whether these factors express themselves in ways that have yet to be explored, ways that could include the disruption of a deceiver's low-level motor control at a computer.

The adaptation of an existing physiological lie-detection technique to a computer-mediated environment might serve as a first step towards identifying these behaviours. This adaptation would need to be based on the assumption that some form of overt limbic affect results from one or more of the psychophysiological correlates of deceit. One such correlate results from the response-time latency effect of the orienting reflex (OR) within a Guilty-Knowledge test. This



phenomenon is of particular interest because investigators have observed overt limbic behaviours after the elicitation of an OR.

Farwell and Donchin (1991) developed a variant of the GKT that used an 'oddball paradigm' to detect the guilty-knowledge of a crime. The test involves a categorisation task that requires a participant to classify a series of two-word phrases into two groups. The phrases (stimuli) would appear to the participant in a randomised sequence and the participant would press one of two buttons depending on the category membership. One set of phrases, the 'target' phrases were learnt before the test and occurred less frequently than the second set of 'irrelevant' phrases. When a participant saw a target phrase they were to press one button, and when they saw any other 'irrelevant' phrase they were to press the other button. As the familiar target phrases occur less frequently than the irrelevant phrases, they are likely to elicit a P300 event-related potential (ERP) on an electroencephalogram (EEG), unlike the irrelevant items. The guilty-knowledge component of the test involved adding certain crime-relevant phrases (the 'probes') into the group of irrelevant phrases. The central thesis of the technique is that a participant who was aware of the crime is more likely to elicit a P300 for the probe items because they are able to distinguish these items as a third, less frequent subcategory.

Farwell and Donchin conducted an experiment using this technique. Twenty university students carried out an espionage-styled mock-crime. There were two different scenarios available and each participant was only aware of one of them. However they received the test twice, once for each scenario and these formed the *Guilty* and *Innocent* within-subject conditions for the experiment. The participants were made to remember six pertinent pieces of information that related to the 'crime' they were going to commit. The next day, they were then sent to meet a person at a predetermined location, exchange a password and receive a 'secret' file. The participants were tested following the successful completion of this task.

As predicted, the participant ERPs for probe items in the Guilty condition resembled that of the target items, whereas the ERPs for the probe items in the Innocent condition resembled that of the irrelevant items. Farwell and Donchin also recorded the response times (RTs) for each participant and found that the mean RTs of the probe items were generally slower than the irrelevant items in the Guilty condition but not in the Innocent condition. It is this RT latency effect that is central to the current study.

Skeleto-muscular effects have been associated with the elicitation of the OR since its discovery. The attempts of Pavlov's students to demonstrate a conditioned response to their Professor were often frustrated by the dog's physical orientation to Pavlov himself (Lynn, 1966). Other than a physical orientation of the senses to the novel stimulus, the OR may also cause a number of other mechanical effects, "ongoing reactions are temporarily arrested and general muscle tonus rises, increasing readiness for activity in the skeletal muscles" (Lynn, 1966; p.3). It is not entirely clear what this effect could have on RTs themselves. The inhibition perspective suggests that the orientation of perceptual apparatus will mutually confound motor responses (Routtenberg, 1971; cited in Spinks and Siddle, 1983). A study by van Olst (1971) showed that

RTs slowed down when an OR was evoked using an auditory stimulus (a change in tonal frequency). Conversely, the facilitation perspective suggests a speeding up of RTs. Blackman (1966) found some evidence to support this claim although initially he found an inhibitory effect. The RT data in Farwell and Donchin's study would also seem to support the inhibition perspective. It would seem as though the elicitation of the OR resulted in a momentary arrest that slowed the participant's button presses. Despite this finding, the authors voiced a reservation about using the RT as a dependent measure for guilty-knowledge as it susceptible to manipulation (countermeasures) by the participant.

This reservation has been challenged by Seymour et al. (2000) who cite research examining the participant manipulation of RTs. They refer to an experiment by Ratcliff and McKoon (1981) that investigated Posner and Snyder's (1975) two-factor theory of recognition. Ratcliff and McKoon had investigated the effects of the automatic and strategic components of word priming using an experimental setup that involved recording RTs through key presses. They found that the strategic function of facilitation (whereby participants were given information in advance about the relationship between the priming word and the potential target word) took longer than 700 ms. Seymour et al. stated that this provided some evidence that strategic control of responses under 800 ms was too difficult and that Farwell and Donchin's reservation would be unwarranted as long as the participants were encouraged to make quick decisions.

Seymour et al. (2000) replicated Farwell and Donchin's experiment in order to examine the effectiveness of RTs is more depth. They conducted three experiments using a similar procedure and the exact same set of stimuli. The first experiment involved a mock 'computer-crime' that was designed to be a bit more realistic than the somewhat theatrical Farwell and Donchin espionage scenario. The participants were then tested in a similar manner, but only RTs were recorded. The second and third experiments mimicked the first but were designed in order to address Farwell and Donchin's reservation concerning countermeasures. In the second

ITEM	CONDITION	
	Guilty Probe	Innocent Probe
	Guilty Irrelevant	Innocent Irrelevant

Figure 6.1: Experimental Conditions used in (Farwell and Donchin, 1991)

experiment, the participants were explained the purpose of the probe, target, and irrelevant items, and in the third they were informed of the expected pattern of results (i.e. slower RTs for the probes). In the first experiment, a repeated measures ANOVA (Condition X Item) revealed a main effect of guilt where probe item RTs were on average 300ms slower in the Guilty condition than in the Innocent one<sup>1</sup>. Similar effects were found in the other two experiments although the differences were slightly smaller. This effect (larger than that observed by Farwell and Donchin) would seem to suggest the potential of RTs to be used as fairly reliable dependant measure of guilty knowledge. A discriminant function analysis yielded an overall discrimination accuracy of 95% in the first experiment and the results were resistant to the countermeasures that were tested in the second and third experiments. Like Farwell and Donchin, the investigators summarised RTs for each participant according to the mean RT for each stimulus type. The consequent ANOVAs were then conducted on these values and no information was reported concerning the intra-subject distribution of RTs that contributed to these means.

An RT measure, reliably inhibited by the OR and detectable through a mouse or keyboard would provide one doorway into a field of real-time CMC lie detection. However, the reliability and validity of the RT is far from proven, and it is important to recognise that although Seymour et al. identified an average difference in RT of around 300 ms, the results of Farwell and Donchin are considerable more conservative (79 ms).

## 6.2 Hypotheses

The purpose of this study is to re-examine the oddball paradigm using a crime scenario that is more realistic and more relevant to the CMC domain. In addition, the attempt will be made to record response times through a standard computer mouse as opposed the more artificial experimental setup used in previous studies. The success of the technique may then provide justification for further investigations into the detection of the reflex responses at the computer-interface and how those reflexes could be used to detect deceit.

This study tests the hypothesis that participants will be significantly slower in probe item RTs when they have guilty knowledge of a crime than when they are innocent of a crime ( $H_1$ ). As well as looking at an effect across participants, the analysis aims to determine whether such a difference can be used to determine guilt for a single participant, using only their 'Innocent' test as a reference ( $H_2$ ).

## 6.3 Method

The experimental procedure was based on the protocols described by Farwell and Donchin (1991) and Seymour et al. (2000). This oddball paradigm involves the participants undertaking a mock crime, and then being tested for the guilty knowledge of that crime and another scenario

---

<sup>1</sup> This comparison required the selection of a set of innocent probe items that were chosen at random. The same set was used across all participants.

with which they had no knowledge. This experiment involved a novel crime scenario whereby the participant created a disturbance within a computer-mediated chat room. The test was conducted at a standard desktop computer terminal using a standard two-button mouse to record response times.

### **6.3.1 Participants**

Thirty participants (17 male, 13 female, mean age = 21 years, SD approx. 1 year, 4 months) were recruited for the study. The participants self-selected through an advertisement posted throughout the University. Eighteen of the participants were undergraduates and two were postgraduates. They were paid £5 on conclusion of the procedure that took approximately one hour.

Five postgraduate volunteers also took part in an activity to generate the simulated chat transcript (described in the next section).

### **6.3.2 Design and Protocol**

The experimental protocol for each participant followed of a sequence of phases that are described below. A fuller description of the main laboratory, software, and preliminary development has been provided in section 6.3.3.

#### **6.3.2.1 Arrival**

The participants were greeted on arrival and shown to the main laboratory. On route to the laboratory, the investigator indicated that the experiment was being coordinated with another group of volunteers who were participating in a concurrent study to investigate group interaction in CMC. This false claim was made in order to reinforce the mock crime scenario.

#### **6.3.2.2 Briefing**

On arrival in the laboratory, the participants were briefed and provided with the 'Participant Information Sheet 1' (see Appendix II). In the briefing, the participants were informed that the purpose of the study was to conduct a 'simple lie-detector test' using a standard home computer.

They were told that in another part of the building, five volunteers, involved in a concurrent study, were about to begin a task to solve a puzzle (a three-dimensional cube that had been broken into six distinct pieces). Each team member had one piece of the puzzle and the purpose of task was to reassemble the cube using only a computer-mediated chat program to communicate and describe the solution. The participant was told that the team members were expecting a sixth person to join the group, and that they were going to join in as the sixth member. However, unlike the other team members, the participant was told that they were to be given a different objective to the others. It was their objective to disrupt the team's discourse in

a particular manner to be described by the investigator. This disruption had been designed to cause sufficient confusion that the team would fail in their task.

They then signed a consent form for the study.

### **6.3.2.3 Mock Crime**

The participant was then presented with information they needed in order to commit the disruption (see Appendix II). They were told that they needed to learn the information, as it would not be provided during the chat session. The crime involved the participant logging into the chat system. After login, they were to tell the other team members about 'their piece' of the puzzle, a fabricated piece that was not a part of the final solution. Following this, the participant had to continuously query another member of the chat room with clarification requests, so as to distract them from the task at hand. They were then to type the phrase "BROKEN MOUSE" into the chat room program repeatedly until it filled up the history screen, thus making it difficult for the other team members to follow the conversation. They were told that they would be given eight minutes to complete the task, at which point they were asked to logout of the chat room system.

The participant was then told that they were to be given five minutes to learn the information needed to commit the disturbance, and that software had been provided to help them to do this. The software (described later) provided three learn-test cycles of six pertinent 'crime' items (their chat room pseudonym, password, the puzzle type, fabricated puzzle piece, team member to query, the phrase 'Broken Mouse'). The investigator left the room whilst the participant learnt the information. The reason given for this departure was to apparently check with the co-investigator that the other team was ready to start the exercise. Again, this fabrication was added in order to enhance credibility.

After five minutes had elapsed, the investigator rejoined the participant. The participant was informed that the other team had started early, but that this was not a problem since they were going to be doing the task for half an hour. This 'unexpected undesired event' was added so as to reinforce the perception that the other (hitherto unseen) chat room members actually existed. The participant was then shown a screenshot of the chat software and then described its operation. The chat software was put into initialisation mode and the participant information sheet was retrieved. The participant was further advised that they were allowed to converse lightly with the other team members. However, they were advised not to participate in any extended discussions in order to reduce the risk of entering into an aggressive conflict with the other team members. The real reason for this directive was to try to restrict the participant's interaction with the system. This was necessary because extensive interaction could draw attention to the limitations of the simulation. After being asked if they had any further questions, the investigator left the room for eight minutes. During that period, the chat-room program provided a simulation of the team exercise that was 'pre-recorded' in a preparatory study (described in section 6.3.3).

After eight minutes had elapsed, the investigator rejoined the participant and asked them how the session went. One purpose of this debrief was for the investigator to probe the effectiveness of the simulation.

#### **6.3.2.4 Distracter task**

The distracter task was designed to prevent rehearsal of the crime information. Computer games are well known for their ability to immerse and engage players, and a simple game written by the investigator was presented to the participant. The game was designed to occupy the participant for ten minutes.

#### **6.3.2.5 Testing Phase**

Each participant was tested twice (corresponding to the Guilty and Innocent conditions respectively). The order in which the tests were presented was counterbalanced across participants. For each test, the participant was asked to memorise a set of six target items with the help of the same piece of training software that was used to remember details of the mock-crime. During the test, these target items were presented to the participant, intermingled with a set of irrelevant items. When the participant identified one of the target items, they were to click the left mouse button. For all other items they were to click the right mouse button.

Each test comprised of target, probe, and irrelevant items. Six target items, six probe items, and twenty-four irrelevant items (four for each of the probe items) were repeated three times and presented to the participant in a random order. The full list of all the items used has been provided in the Appendix. For the Guilty condition, the probe items corresponded to the six items learnt for the mock crime (the 'guilty' knowledge). For the Innocent condition, the probe items comprised of six items that were different, but of a similar type to the crime items. Half the participants took the 'innocent' test first, and the other half took the 'guilty' test first. Before the tests, the participant was shown the software and was given a short practice test. The participant was then given five minutes to learn the target phrases before each test.

#### **6.3.2.6 Debrief**

Following the final test, the investigator re-entered the lab and asked the participant for feedback about the session. The experimental deception was then revealed. Firstly, the investigator asked the participant how they felt about the chat session. Then the revelation was presented in the following manner.

*"It [the chat session] was a very strange situation wasn't it? First of all, you were unable to interact with the other members from the start of task, and also I gave you very specific instructions on what you were allowed to do in the chat room, essentially 'binding your hands' with respect to how you could interact with the others. There was a reason I did this. If I had let you interact, as you would have*

*done normally, then very, very quickly you would have realised that, that was not a real chat room program, and that was in fact, a simulation. There are no 'five people' in another part of the building, because I needed to chat room situation to be exactly the same for everyone taking part in the experiment. Is that okay, how do you feel about that?"*

This revelation was designed to reinforce the idea that the participant was not deceived because of any gullibility on their part. Following this, the investigator explained a little more about the purpose of each task performed throughout the session. The investigator then answered any further questions about the study and asked the participant if there were any ways in which the chat simulation could be made more convincing. The point was reiterated that the participant was still free to withdraw and this was followed by payment.

### **6.3.3 Preliminary Work and Apparatus**

The procedure reported in section 6.3.2 demanded the production of four pieces of bespoke software and the production of a chat transcript to be played back during the simulation. These were all developed by the author and are described here in more detail. Screenshots of all the software described in this section have been provided in Appendix II.

#### **6.3.3.1 The Training Program**

The training software was designed to provide the participant with an iterative learn-test cycle to help them commit the mock-crime and target items to memory. The system was written to work under the Java 2 Runtime Environment through Windows 2000. The software would first present the participant with instructions about how to navigate through the test. Upon selecting to proceed, the participant would be presented with the six items that were to be remembered. After the participant had read through the items, they were then presented with a new page with blank fields corresponding to each of the items. Once the participant had completed the fields, the system highlighted any errors and then repeated the learn-test cycle a further two times.

#### **6.3.3.2 The Chat Room Simulation**

The chat room simulation was designed in order to provide a convincing, typical chat-room interface. The system was written to work under the Java 2 Runtime Environment through Windows 2000. The key function of the system was the ability to replay a pre-recorded chat transcript in the history window that would create the impression that the participant was actually involved in a chat room.

On loading the system, the chat system remains in the initialisation state until the participant selects to proceed. At this point, the full interface is presented to the participant and the system begins to play back the chat transcript at one-and-a-half times the speed that it was recorded (see section 6.3.3.6 for a justification of this factor). The participant's 'chat entry' text field

remains disabled until they login by selecting the respective button. The login dialogue prompts the participant for a case-sensitive username and password. If these details are correctly entered the chat entry field is enabled and the username is added to the 'active user' frame.

The participant is then free to conduct the disturbance within the simulation. Utterances can be posted to the chat system using the 'chat entry' field in a similar manner that is used within a typical quasi-synchronous chat program. In order to enhance the perception of interactivity, the system also performed some simple pattern-matching tests on the participant's postings. Upon the entry of any of the crime-specific keywords (such as the piece type, or the expression 'Broken Mouse') the system would insert some pre-fabricated responses. The pre-fabricated responses were posted in intervals of ten seconds after the relevant pattern was identified, until they either ran out, or the next pattern was identified. For example, one of the pre-fabricated responses to the case-insensitive entry of either of the words 'Broken' or 'Mouse' was the chat-room member Jane's casual response of "Call your investigator bloke person". The responses were designed to be informal and included common spelling errors and colloquial language. Naturally, the inserted pre-fabricated responses would momentarily suppress the utterances from the respective team member in the chat transcript (otherwise two utterances might be posted too closely together to create an impression of a sensible typing speed).

Upon conclusion of the mock crime the participant could then select the logout button. At this point, the transcript of the session was saved to disk and the program terminated.

### **6.3.3.3 The Distracter Task**

The distracter task was designed as a 'non-immersive'<sup>2</sup> virtual-reality computer game. The system was developed using the rapid game development environment 'Dark Basic'. The goal of the game was to try and identify the location of five cubic signs hidden around a three-dimensional tropical island. The participant was allowed free-movement around the island and was asked to write down (1) what was written on each sign and (2) the rough co-ordinates of where they had found the sign. The participant was informed that there were a number of cheetahs that were also on the island running about and trying to catch them. If the cheetahs managed to do this then the task (which was being timed) would be over. In actual fact the cheetahs simply ran about at random and the task would never terminate until the allotted time limit of ten minutes had been exceeded. The presence of the cheetahs was employed in order to keep the participant concentrating on the game. In case the participant finished the task early, they were asked to try and determine how many cheetahs existed on the island. The game incorporated ambient sound effects so as to enhance immersion and help prevent the rehearsal of crime-relevant information.

---

<sup>2</sup> This expression is commonly used to describe the virtual environment as being interacted with through a monitor and keyboard, as apposed to suggesting that the participants did not find the task cognitively engaging.



#### **6.3.3.4 The Testing Software**

The testing software was developed to run under the Java 2 Runtime Environment through Windows 2000. Upon initialisation, the testing software read-in the relevant item list and then sat in a wait state until the user clicked the left mouse button. The items were presented on the screen one after the other in intervals randomly varying between 800 and 1200 milliseconds. This randomisation was used in order to prevent rhythmic responses. After an item had been presented to the participant, the system waited for a mouse button click. Upon the click, the next item was presented after the randomised interval period had elapsed. If the participant took longer than one second to respond, the system displayed the message 'Too Slow' before the next item was displayed. Upon conclusion of the test, the system saved all logged mouse-events to file.

It should be noted that the functioning of this testing software mirrors the relevant functions of the COGSYS system described in (Ratcliff and Layton, 1981). Unlike COGSYS, the system was designed specifically for the requirements of this study and was not a hard real-time system. The implications of this latter drawback was deemed acceptable since the predicted error margin of mouse event timestamps would cause minimal noise for an effect size of the order of 100 milliseconds.

#### **6.3.3.5 The Laboratory and Hardware**

The laboratory contained a desk, two chairs, a standard Windows 2000 desktop PC (1.2 GHz Athlon, 128 Mb SDRAM), and a 17" Monitor (resolution 1024x768, 75Hz, 32 bit colour) with a two-button mouse and keyboard. Headphones were provided for audio during the distracter task. Earmuffs were provided for the learning sessions. The participant was also provided with a clay model of the complete block puzzle (illustrated in Figure 6.2).

#### **6.3.3.6 Generation of the Chat Transcript**

In order to generate a convincing transcript for use with the simulation, the investigator actually conducted the block puzzle exercise with five volunteer postgraduates. Each volunteer was given a piece of the block puzzle shown below.

The volunteers were asked to see if they could assemble the block puzzle verbally, using a chat room program as their only means of communication (see Appendix II for instructions given to volunteers along with a sample of the final recorded transcript). The chat room software used was the same one developed by the author for the studies described in the next two chapters. This program recorded the entire transcript of the exercise along with the timing information of when each utterance was posted. The transcript could then be given directly to the chat room simulator after all the pseudonyms used by the volunteers were replaced by fictitious names. In addition, the initial introductions were stripped from the transcript so that the simulation could start from the middle of the exercise.

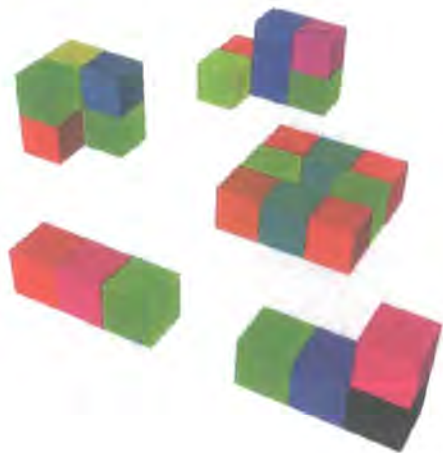


Figure 6.2: The Block Puzzle

In pretests, it was found that this recorded simulation was too slow for the participant to feel involved. This could have been because the participant was not fully immersed in solving the puzzle. A sense of interactivity was deemed a necessary requirement of the simulation in order to enhance realism. It was found that by speeding the playback of the transcript by one and a half times, this sense of interactivity was enhanced.

6.4 Results

The debrief revealed that the experimental deception had been successful. Only one participant indicated some suspicion concerning whether or not the other 'members' were collaborating with the investigator. Reports varied in the degree of interactivity that participants felt that they experienced. Interestingly, more experienced chat room users perceived more interaction from the other members than the less experienced users. This is consistent with the predictions of SIP whereby more experienced CMC users are more likely to have heightened impressions of interactivity based upon their experience with this mode of communication.

All participants were able to successfully learn and execute the crime scenario. None of the participants withdrew or showed discomfort at the experimental deception. Unless stated otherwise, all statistical tests are reported with  $\alpha = 0.05$ .

6.4.1 Accuracy

The accuracy of the responses has been provided in Table 6.1. Accuracy was assessed as the percentage of correct mouse button clicks. From Table 6.1, it can be seen that the guilty probes have a lower accuracy and a greater spread than the innocent probes, resembling the target items. A 2 (Guilty:Innocent) X 2 (Probe:Irrelevant) repeated measures ANOVA was conducted on the data and this revealed a significant main effect of guilt ( $F_{1,29} = 16.25, p < 0.001$ ), a main

effect of stimulus type ( $F_{1,29} = 17.33, p < 0.001$ ), and a guilt by stimulus type interaction ( $F_{1,29} = 20.06, p < 0.001$ ). The chart in Figure 6.3 shows that these effects are heavily influenced by the comparatively low accuracy of the guilty probe responses. The probe items are also less accurate than the irrelevant items within the guilty test, but the reverse is true within the innocent test (however the difference is not very large).

Condition	Stimulus	Mean (%)	Std. Deviation
Guilty	Target	79	12
	Probe	90	12
	Irrelevant	99	1
Innocent	Target	82	11
	Probe	99	2
	Irrelevant	99	1

Table 6.1: Accuracy (%) of Responses Across Participants (N = 30)

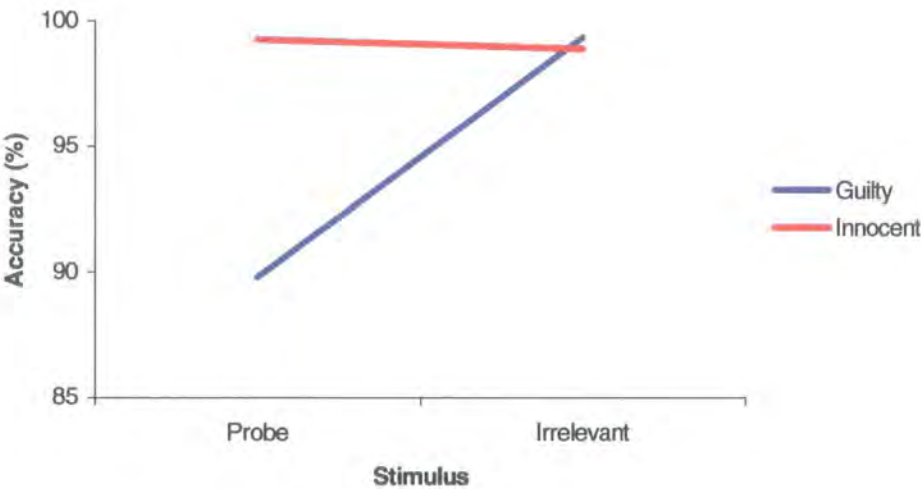


Figure 6. 3: Accuracy (%) of Responses

The lower accuracy of the guilty probe items has a further implication for the assessment of guilty knowledge. It could be surmised that the reduced accuracy of probe items within the guilty condition is in itself an indicator of guilt and this finding is discussed further in section 6.5.

6.4.2 Overall Comparison of RTs

In order to assess the overall hypothesis that probe items were significantly slower than irrelevant items within the guilty condition, a repeated-measures ANOVA was conducted across all of the participant's mean response times for each stimulus type. The means for each participant have been provided in Table 6.3. The ANOVA revealed a significant main effect of

guilt ( $F_{1,29} = 66.93, p < 0.0001$ ), a main effect of stimulus type ( $F_{1,29} = 42.97, p < 0.0001$ ), and a guilt by stimulus type interaction ( $F_{1,29} = 46.48, p < 0.0001$ ).

Condition	Stimulus	Mean (ms)	Std. Deviation
Guilty	Probe	612	96
	Irrelevant	519	63
Innocent	Probe	481	69
	Irrelevant	479	54

Table 6.2: Correct-Response Times (ms) Across Participants (N=30)

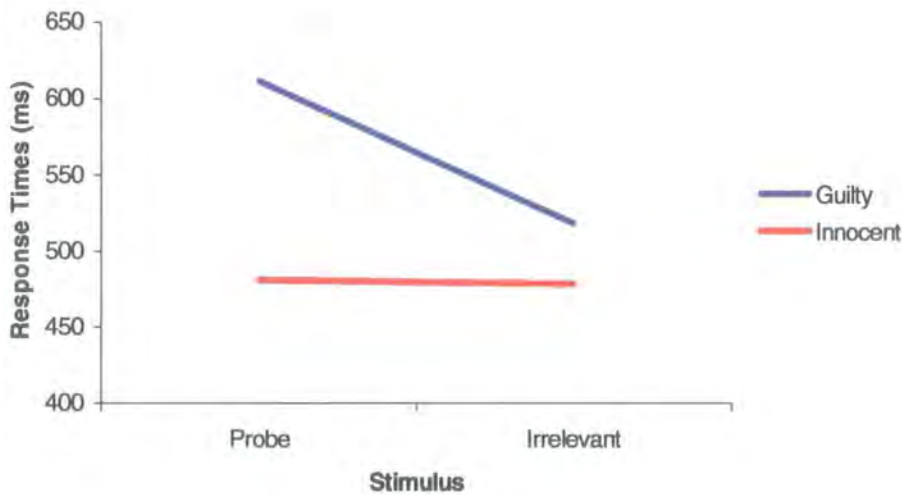


Figure 6.4: Correct-Response Times (ms)

It can be seen from Figure 6.4 that there is an interaction that would suggest that probe items are significantly slower than irrelevant items in the guilty condition, but not so in the innocent condition. However, there is also a main effect for guilt that is not entirely attributable to the guilty probes. Guilty irrelevant responses were significantly slower than those responses attained by innocent irrelevant responses and box plots of these two RT have been provided in Figure 6.5. The box plots indicate that although there is a significant difference, the magnitude of this difference is not very large (38ms between means), and the distributions between the innocent and guilty irrelevant items are similar.



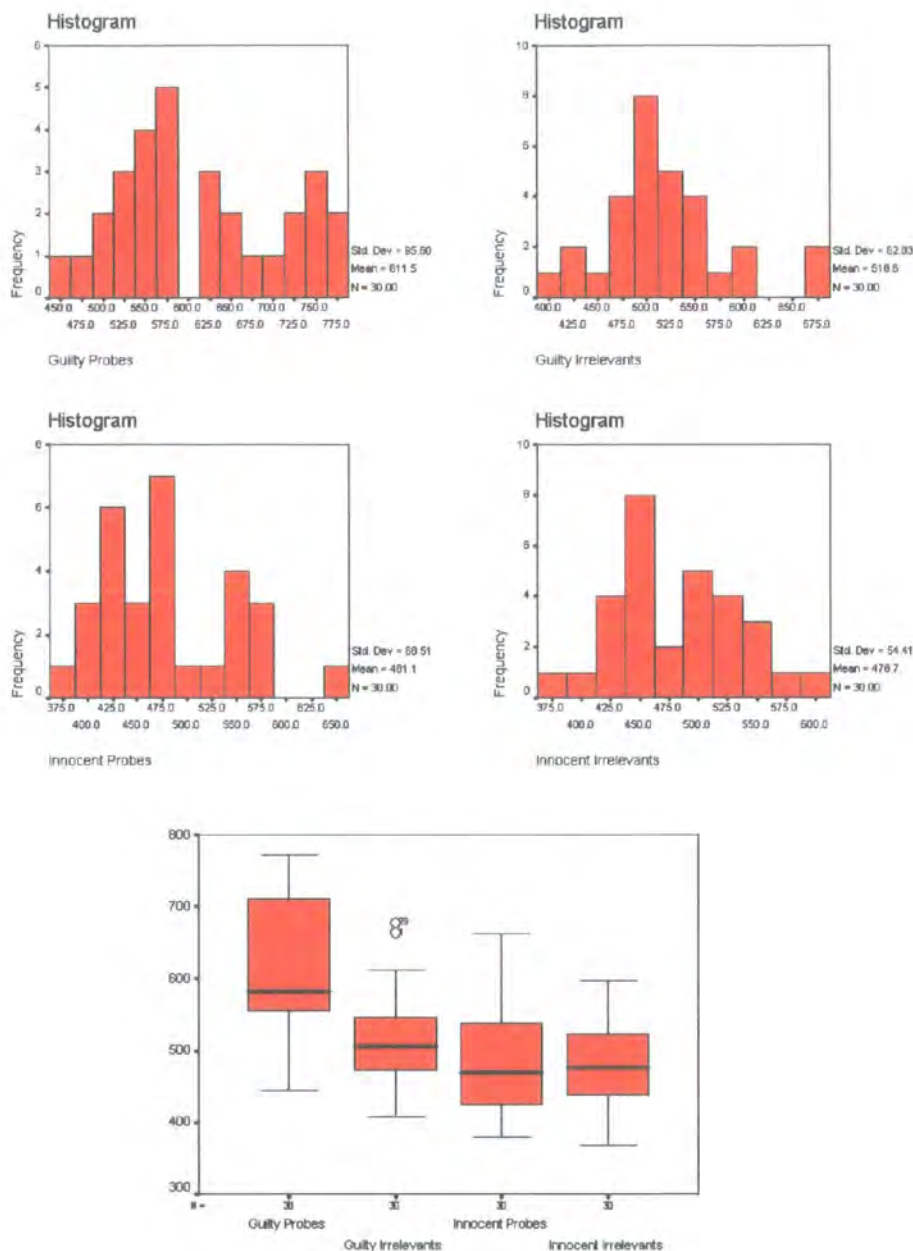


Figure 6.5: Box-Plots and Distributions of RTs

### 6.4.3 Individual Comparison of RTs

Seymour et al. (2000) employed a technique that could be used to assess the guilt status of an individual participant. The technique assessed whether the guilty probe responses are from the same population as the guilty-irrelevant, innocent-probe, and innocent-irrelevant RTs. These comprised of three tests whereby a low significance on any specific test would indicate that the two response blocks are from a different population, signifying guilt. The three tests include:

- 1. A Fisher exact (2X2) test with a one-tailed hypothesis predicting a higher guilty-probe error rate
- 2. A Kolmogorov-Smirnov two-sample test with a one-tailed hypothesis predicting a longer distribution for the guilty-probes
- 3. An F-test for variances with a one-tailed hypothesis predicting a higher variance for the guilty-probes.

The Fisher exact test compares the guilty-probe error rate against the innocent-probe error rate. The other tests compare the guilty-probe RTs with the pooled RTs from the guilty-irrelevant, innocent-probe, and innocent-irrelevant responses. The significance results from these tests have been presented in Table 6.4, and the mean RTs have been presented in Table 6.3.

Probe		Pool				Probe		Pool				Probe		Pool	
	M	SD	M	SD		M	SD	M	SD		M	SD	M	SD	
1	738	180	609	176	11	556	230	411	105	21	744	279	608	223	
2	577	123	514	105	12	558	124	457	117	22	566	222	463	131	
3	639	154	489	121	13	772	262	510	142	23	726	205	452	264	
4	659	273	487	151	14	613	201	506	148	24	629	116	447	118	
5	631	334	480	136	15	444	143	387	121	25	582	165	478	146	
6	519	188	453	125	16	571	195	538	139	26	513	128	451	128	
7	710	172	510	136	17	463	106	523	135	27	561	181	500	98	
8	719	240	523	146	18	528	272	493	220	28	499	106	419	92	
9	765	136	573	106	19	581	94	547	165	29	755	202	601	184	
10	495	106	471	116	20	674	310	520	200	30	558	166	466	131	

Table 6.3: Mean Response Times

Fisher				K & S				F				Fisher				K & S				F			
1	1.000	0.004	0.407	11	0.500	0.008	< 0.001	21	0.146	0.007	0.104	11	0.500	0.008	< 0.001	21	0.146	0.007	0.104	11	0.500	0.008	< 0.001
2	0.114	0.086	0.176	12	1.000	< 0.001	0.335	22	0.243	0.044	0.001	12	1.000	< 0.001	0.335	22	0.243	0.044	0.001	12	1.000	< 0.001	0.335
3	0.500	< 0.001	0.073	13	0.010	< 0.001	< 0.001	23	0.004	< 0.001	*	13	0.010	< 0.001	< 0.001	23	0.004	< 0.001	*	13	0.010	< 0.001	< 0.001
4	0.243	0.001	< 0.001	14	1.000	0.003	0.027	24	0.243	< 0.001	*	14	1.000	0.003	0.027	24	0.243	< 0.001	*	14	1.000	0.003	0.027
5	0.243	0.031	< 0.001	15	0.243	0.087	0.152	25	1.000	0.002	0.216	15	0.243	0.087	0.152	25	1.000	0.002	0.216	15	0.243	0.087	0.152
6	0.500	0.058	0.005	16	0.243	0.480	0.022	26	0.500	0.073	*	16	0.243	0.480	0.022	26	0.500	0.073	*	16	0.243	0.480	0.022
7	0.008	< 0.001	0.117	17	1.000	*	*	27	1.000	0.137	< 0.001	17	1.000	*	*	27	1.000	0.137	< 0.001	17	1.000	*	*
8	0.500	< 0.001	0.001	18	0.052	*	0.113	28	0.514	0.002	0.176	18	0.052	*	0.113	28	0.514	0.002	0.176	18	0.052	*	0.113
9	0.023	< 0.001	0.087	19	1.000	0.060	*	29	1.000	< 0.001	0.268	19	1.000	0.060	*	29	1.000	< 0.001	0.268	19	1.000	0.060	*
10	1.000	0.265	*	20	*	0.038	0.002	30	1.000	0.014	0.071	20	*	0.038	0.002	30	1.000	0.014	0.071	20	*	0.038	0.002

\* Indicates an effect that acts contrary to the tailed-hypothesis

Table 6.4: A Comparison between Guilty Probes and Pooled Data

According to the scheme proposed by Seymour et al. (2000) the guilty tests were correctly classified 76.7% of the time at  $\alpha = 0.05$ , and 73.3% of the time at  $\alpha = 0.01$ . The false-negatives consisted of the guilty tests for participants 2, 10, 15, 17, 18, 19, and 26. The only test that showed a guilty-probe mean that was quicker than the pooled items was the test conducted with participant 17.

An estimation of the false-positive rate was calculated by conducting the Kolmogorov-Smirnov two-sample test and the F-Test for variances on the innocent-probe items against the pooled innocent-irrelevant and guilty-irrelevant items<sup>3</sup>. Guilty-probe RTs were omitted from this test and so no Fisher (2X2) test was conducted, however, this did not affect the results, as the innocent-probes were all highly accurate. These results are shown in Table 6.5.

K & S			F			K & S			F			K & S			F		
1	*	*				11	*	*				21	0.064	0.1606			
2	0.0065	*				12	*	< 0.001				22	*	*			
3	*	*				13	*	*				23	*	*			
4	0.4695	*				14	*	*				24	0.288	*			
5	*	*				15	*	0.002				25	0.172	*			
6	*	0.4773				16	0.33	*				26	*	0.288			
7	*	*				17	0.101	*				27	*	*			
8	0.001	*				18	*	*				28	*	0.009			
9	*	*				19	*	*				29	*	*			
10	0.35	*				20	0.053	*				30	*	*			

\* Indicates an effect that acts contrary to the tailed-hypothesis

Table 6.5: A Comparison between Innocent Probes and Pooled Data

The table indicates that approximately 83.3% of innocent tests were classified correctly at  $\alpha = 0.05$  and  $\alpha = 0.01$ , leaving only an estimated 16.7% false-positive rate. The false-positives consisted of the guilty tests for participants 2, 8, 12, 15, and 28.

### 6.5 Summary and Discussion

The results confirm the hypothesis that participants RTs are slower when they react to guilty knowledge items within the test. Further it was found that these results could be obtained with a standard two-button mouse using a standard PC. Seymour et al.'s (2000) analysis technique correctly assessed 76.7% of the guilty tests at the 95% chance level and 83.3% of the innocent tests. Nonetheless, key differences were found between the data obtained in this experiment and that obtained by Seymour et al. (2000).

<sup>3</sup> Seymour et al. (2000) only considered innocent-irrelevant items when assessing their false-positive rate.

The responses recorded in this study were on the whole more accurate than they were in the three experiments reported by Seymour et al. (2000). Responses for the innocent items and the guilty-irrelevant items had an accuracy level of approximately 99%, and guilty-probes were rated at 90%. Seymour et al. (2000) reported a high of 87% for the irrelevant items and 81% for the guilty probes. More significantly, their accuracy ratings for the innocent probes show values of 96%, 99%, and 99% for each of their three experiments. This means that there was a much larger difference between the accuracy of the guilty and innocent probes in the Seymour et al. study than there was in the current study. This is an important difference because it would affect the outcome of the Fisher (2X2) test. Seymour et al. (2000) report the correct classification of the guilty tests at 98% at the 95% chance level, as opposed to the more conservative 76.7% attained by this study. It is possible that the lower level of accuracy attained by the guilty-probes in the Seymour et al. study contributed significantly to this difference.

Response-times were also considerably faster. Seymour et al. (2000) report the mean RTs for guilty-probes in their first experiment as exceeding 900ms, in this study they were reported at 612ms. Their Innocent items were also recorded at around 600ms, which again are slower than the 480ms figure that was attained for this study. The difference in RTs between the guilty and innocent items was also more conservative. In the Seymour et al. study, the guilty RTs were approximately 300ms slower than the innocent RTs and the results achieved here suggest a figure less than half that size (roughly a difference of just over one standard deviation).

There was also an unexpected interaction for guilt with the irrelevant items, which showed slower RTs in the guilty condition. It seems probable that the presence of the guilty-probes acted to make the participants more cautious, resulting in slower button presses. This interaction mean that the between-subjects analysis used above, and employed by Seymour et al. could in fact be too conservative (although not to the extent where the RT difference approaches 300ms). In addition it means that it may be inappropriate to pool guilty-irrelevant RTs with the innocent data in the individual analysis as these come from essentially different populations. It might then be more appropriate to conduct the three tests between the guilty and innocent blocks of an individual examinee.

The results reported in the current study are considerably more conservative than those reported by Seymour et al. (2000) and are consistent with the difference in RTs observed by Farwell and Donchin (1991) in their original experiment. Nonetheless the experiment did reveal that reflex actions that reveal guilty knowledge could be detected using standard two-button mouse on an ordinary PC. However, it is possible that use of the mouse allowed responses that were too fast and that this mediated the guilty knowledge effect. In addition, it was found that the use of a novel form of realistic computer crime could be achieved through an effective simulation, and that participants that conduct this form of online episode are also susceptible to conventional lie-detection techniques.

It seems as though there is now some evidence to suggest that physiological leakage cues are detectable at a low-level through the computer interface. However, the extraction of such cues



is likely to be restricted to situations where interface control is relatively uninhibited and continuous. This level of interaction is typical of a user's behaviour when typing at a keyboard, particularly if they are deeply immersed within a chat room. A keystroke analysis of this behaviour might reveal affective clues that correlate with deceit. Orienting behaviour may be detectable as well, possibly being less controlled behind the veil of CMC. Such behaviours could result in intermittent pauses at certain points within the discourse or longer latency periods between keystrokes. This sort of application may be useful for naturalistic lie-detection, but it is probably of more use within the context of a CMC-based investigative interview where the discourse is kept under a greater level of control. In this situation, such an analysis could accompany a CBCA or RM rating. The GKT could indeed be conducted entirely using a piece of CMC chat software or could be incorporated as part of a wider interview. These are possibilities for future research.

## Chapter 7

# Investigative Interviewing with QS-CMC

A strategy for CMC-based investigative interviewing, based upon the Enhanced Cognitive Interview (ECI) is presented. The technique was evaluated by using discourse-analytic techniques within a quasi-experimental scenario. Participants were shown a video clip and the adapted ECI was used to interview them about what they could remember. The interview was conducted over a QS-CMC chat program that was written by the author. The transcripts were analysed in a qualitative and explorative fashion that focused upon conversation management during the interviews. Evidence was found that the adapted ECI could be used for SVA because it resulted in detailed statements with low amounts of confabulation. Whilst these results are carefully qualified, it was also found that the discourse within the interviews did not suffer the detrimental effects posited by early CMC theorists, and that the adaptive processes posited by contemporary theories, facilitated the interaction. Further suggestions were made that could improve the adapted ECI in preparation for the study described in Chapter 8.

---

### 7.1 Background

Lie-detection is difficult. The literature reviewed in Chapter 2 showed that the very concepts of deceit and lying are fluid and subjective, motivations are difficult to ascertain, and the distinction between intent and mistake are rarely clear-cut. That is why many lie-detection techniques are incumbent on not only defining a deceptive context but on actually manipulating that context directly. Controlled lie-detection forces a deceiver to have to lie in a particular manner; a manner that facilitates detection. This form of lie-detection is freed from the difficulties of scrutinising naturalistic lying, where lies and truth are interwoven, and interpersonal reciprocity would seem to modify behaviour over and above deceptive clues (c.f. Interpersonal Deception Theory).

As with an experiment, the greater degree of contextual control, the more validity can be ascribed to the results. Within SVA, the context is manipulated by means of the investigative interview. In some SVA studies, the interviews have not imposed sufficient control, and in some others they have controlled in the wrong way. 'Control' within the context of SVA concerns the effective elicitation of an appropriate level of unconstrained and uncontaminated narrative. Controlling the context does not imply the strict regulation of action, as would be the case in a polygraph test, and so perhaps it is more appropriate to refer to *focusing* on context. If the interviewing style is out of focus, then the cues to deceit or veracity will be inappropriate. The Validity Checklist is a way of assessing contextual focus, whilst techniques such as the Cognitive Interview are a means of attaining that focus. Contextual focus within SVA may be lost in a number of ways and one way might result from the effect of nonverbal contagion. The

degree of focus variability is a key factor in determining an appropriate deceptive context for a lie-detection technique. Therefore by carrying-out a CMC-based interview within SVA, variability could be reduced by occluding nonverbal cues and by focusing the interviewee on the content of their statements. Such a modification may represent a significant qualitative departure from the face-to-face domain and so CBCA might need to be modified according to the peculiarities of the medium. Practically, the technique could not be used with eyewitnesses who have a low level of computer literacy. However, at some expense of applicability the CMC method could evolve as a particular flavour of SVA that affords a high level of contextual focus and therefore a greater probability of validity.

This chapter explores how a CMC-based investigative interview might be constructed along the framework of the Enhanced Cognitive Interview. To this effect, a set of guidelines for the interviewing style has been derived from the existing literature and a quasi-experiment conducted for the purposes of evaluating and refining the technique. Much of the relevant literature on the principles of effective SVA interviewing and the structural nature of QS-CMC was covered in Chapters 3 and 4. This will be supplemented by a review of the research that examines CMC interviewing as a research tool. This literature will form the basis of the derivation provided in section 7.2.

### **7.1.1 Interviewing through CMC**

To date, very few studies have evaluated the merits of synchronous or quasi-synchronous CMC interviewing. In fact, there has been very little empirical work that has even employed the techniques. Although this probably reflects the novelty of the technique rather than anything else, the method's application and scope still awaits a more substantive analysis. This section will review what little information can be gleaned from previous studies under the headings of (1) interaction management, (2) questioning, (3) interview length, (4) anxiety, and (5) technical facilitation. In addition, this section will begin with the presentation of a checklist for task-oriented online interaction that was produced by Mann and Stewart (2000).

#### **7.1.1.1 A Checklist**

In a 1995 article, Nancy Baym proposed five variables that represent the main sources of influence on task-oriented online interaction. In a chapter on 'Online Interviewing', Mann and Stewart (2000) derived a CMC interviewing checklist from these five variables. This checklist is comprised of the following questions:

1. *What is the purpose of the interaction?*
2. *What is the temporal structure of the research?*
3. *What possibilities and limitations for interaction are built into the software?*
4. *What are the characteristics of the interviewer and the participant?*
5. *What is the external context of the research?*

The purpose of the interaction within an SVA interview is the elicitation of episodic memory and the external context of the research is forensic lie-detection. The temporal structure of the interaction is that of a single quasi-synchronous interview, and so the focal point of this study will be on the third and fourth items in this checklist. Under the ecological paradigm, it is impossible to refer to the affordances of technology without referring to the individual actors that use the technology. Therefore within this study, the goal will be to explore the affordances available to an interviewer and an interviewee within a standard CMC chat-system. This study will also explore the affordances that might be made available if the software was enhanced or otherwise modified.

#### ***7.1.1.2 Interaction Management and Rapport***

CMC Interviewers have reported that the management of even the most basic interactional elements of their interviews can be difficult. Annette Markham (1998) carried-out a number of interviews with pseudonymous participants over the Internet. She writes that at first:

“I found it difficult to manage the basic elements of conversation, such as taking turns at the appropriate time, nodding, or mm-hmming to imply ‘Go on, I’m listening’ “

These difficulties highlight the important role of nonverbal cues in the coordination of conversation. Markham mentions a few consequences of the occlusion of these cues including a tendency to race ahead and not allowing enough time for the interviewee to answer properly. She also found herself interrupting the interviewee often, feeling “compelled to fill the blank, black void with more green writing”. Contrastingly, Chen and Hinton (1999) report that the online interviewer seems *less* able to interrupt and that this subsequently limits their control. O’Connor and Madge (2001) conjecture that because CMC represents a ‘freedom from the body’ and because interaction is carefully managed through words, then this can result in a more sociable and friendly discourse. However, their evidence for this was surmised from interviews with new-mothers whereby rapport was augmented with external non-CMC based interaction (such as through personal web pages).

Mann and Stewart (2000) wonder to what extent it is actually possible to establish this mental and emotional connection with the interviewee. They mention that listening skills need to be reengineered within the CMC environment and that “online listening needs to be expressed as words, not silence”. Such expressions may include prompt replies to the interviewee as well as the frequent delivery of verbal assurances.

It seems as though the occlusion of conventional co-ordination cues may confound the constitutive management of structured interaction. The extent to which this occurs remains unclear and this alone would not prevent the successful establishment of rapport so long as speakers were able to sufficiently adapt to the medium (c.f. SIP). In addition, the correct

interpretation of nonverbal cues within any form of interaction is precarious at best, and may be a source of impressional confound, as well as a source of regulatory feedback and control.

#### **7.1.1.3 Questioning**

The principal goal of a question is to receive a full answer, but some authors report that the CMC environment promotes a quick paced interaction that could run contrary to this aim. O'Connor and Madge (2001) conducted sixteen interviews in which they noticed that power relations were in some part decided by typing speed and the use of textual paralanguage. In an interview where power is contested, this could lead to a reduction in answer length and quality. Chen and Hinton (1999) report that in their study "the interviewees felt rushed to respond" and they attribute this to a lack of reassurance of the interviewer's attentiveness. Chen and Hinton also report that their interviewees felt compelled to type before they thought, but that they still managed to produce structured and ordered transcripts.

Language use in questioning still remains an important issue in CMC. Uniquely, the use of ideographs and other similar forms of textual paralanguage need to be used with care. Mann and Stewart (2001) suggest that some participants may not see this as professional research conduct, and this is naturally a feature to be considered under item four of their checklist. O'Connor and Madge (2001) also found that they could run into difficulties when wording responses.

"There were occasions when we were 'lost for words', taking some time to decide on our appropriate response, because we felt like our written comments sounded banal or our questions too direct and leading"

Naturally, leading questions would be a detriment to the ECI. This could be minimised by constructing a repository of question stems that could be copied into the CMC program and fashioned into the discussion (Chen and Hinton, 1999).

#### **7.1.1.4 Interview Length**

Authors have reported that online interviews can take longer than offline ones (Chen and Hinton, 1999; Markham, 1998). This is not surprising because it takes longer for someone to type than it does to talk. Markham points-out that the online interviewer needs to have patience, and should be concerned about responding too quickly (at one point she reports sitting on her hands in order to prevent herself from typing).

The effects of a longer interview length may be tempered by the reports that CMC interviews are easier to keep on track. Specifically, O'Connor and Madge (2001) report that it was easier for them to adhere to a schedule and to ignore sidetracking effectively in the CMC environment although the reason for this was not addressed.

#### **7.1.1.5 Anxiety**

Chen and Hinton (1999) report that their interviewees felt less nervous about the interviewing process when they were sat in front of a computer. They also report that the interviewees felt physically and emotionally distant and that some even had suspicions of being under surveillance. From the interviewer's perspective, Markham (1998) reports her anxiety surrounding the receipt or attention to her questions and remarks.

#### **7.1.1.6 Technical Facilitation**

Computer-based interviewing allows the leverage of a wide variety of software to help in the interviewing process, remedying those difficulties presented by the medium and augmented the process beyond that which can be achieved face-to-face. Apart from tools that help in the analysis of the interview transcripts (including the ability to record the transcript automatically), Chen and Hinton (1999) mention the usefulness of question-management software. These utilities provide sample questions or question-stems that can be copied into a chat client, reducing the amount of time the interviewer spends constructing their utterances, economising cognitive load and subsequently increasing the interviewer's control.

#### **7.1.1.7 Summary**

An effective CMC-based investigative interview is central to the adaptation of SVA proposed within this chapter. In this study, the enhanced cognitive interview has been chosen as the model on which to base such a technique. The findings reviewed in this section can be summarised as follows:

1. *Interruptions can occur accidentally due to misalignment.*
2. *The interviewee may find interrupting the interviewer difficult because of the problems of quickly asserting control.*
3. *Online listening needs to be performed verbally.*
4. *Interviewees may feel rushed to respond.*
5. *Paralanguage may not be suitable, depending upon the interviewee.*
6. *The interviewees may feel anxious about using the software.*
7. *Online interviews are longer but easier to schedule.*
8. *Interviewer utterance construction can be difficult because extra effort is needed to prevent sounding banal or blunt.*
9. *Specialist software can be used that facilitates the interview process and reduces the pressure on the interviewer*

## 7.2 Aims of the Study

These findings may all be relevant to the adaptation of the ECI, however the relevance of these factors within the context of an investigative interview has yet to be determined. The key objectives of this study are:

1. *To derive an interviewing strategy that is based upon the ECI and that can be used within a CMC environment. To state some key topics for further exploration.*
2. *To pilot this interviewing strategy and to explore these topics, suggesting ways in which the technique might be adapted*
3. *To evaluate the technique's sufficiency as an appropriate interviewing strategy for the purposes of a CMC-based SVA.*

## 7.3 Deriving an Interviewing Strategy

The following guidelines and comments have been derived from the literature presented in the previous section and in earlier chapters. As part of this derivation, a number of key themes have been selected upon which the analysis will be based. These explorative themes have been listed in section 7.3.8 following a description of each of the phases used in the adapted ECI.

### 7.3.1 Phase 1: Greeting and Rapport Building

The initial greeting shall be carried-out offline. This has a practical advantage given the nature of the protocol, but it also has a theoretical one. Participant anxiety can affect memory retrieval, can stifle successful interaction and can consequently lead to an increase in false-negatives. Anxiety can occur due to technological, procedural, and aptitudinal factors.

Although the participants will have a basic familiarity with CMC, they may still feel some anxiety about the bespoke system developed specifically for this study. Jacob Nielsen (1993), an expert in Usability Engineering, has reported that in some software usability tests, participants have even been known to break down and cry in response to a tremendous internal pressure to perform. Therefore, in order to lessen any worries that the participant could have, the interviewer should adopt the following guidelines.

- *Inform the interviewee that they are not being tested with regard to their IT competence and make the purpose of the study clear.*
- *Reduce the formidability of the interviewee's environment and computer.*
- *Explain the automatic transcription facility.*
- *Explain about security and confidentiality.*
- *Explain that the software may fail from time to time given its experimental nature and that the interviewee must not be concerned if this happens.*

- *Explain that the computer is resilient to the user and cannot be damaged permanently (barring malicious attack).*
- *Ensure all of the interviewee's questions are addressed, and that the participant feels at ease to ask those questions.*

Adapted from recommendations in (Nielsen, 1993)

The software developed for this study was tested so that it would be as robust as was deemed necessary. Nonetheless the interviewee shall not be misled with regard to the security aspects of the system, especially since the experimenter will have to concede that the software may fail from time-to-time. As the interview material will not be personally identifying and the interviewee is represented solely by pseudonym, this level of security should prove acceptable.

Anxiety can also occur in response to the procedure itself. This is sometimes the case in experimental and quasi-experimental studies where the setting is unnatural and strange. Similarly aptitudinal anxiety may arise out of the perception that the interviewee is in some way being tested. Bringing the initial greeting and introduction offline will help to humanise the procedure and make it seem less like the interviewee is coldly being subjected to a set of dehumanising tasks. The introduction will be performed more efficiently offline and will provide some initial impression information concerning the affective state of the interviewee. This information will be a useful indication of the degree to which the interviewee will need further reassurance. This offline meeting could result in salient first-impression biases, and the interviewer should be aware of this.

The rapport building process (the second part of Phase 1) is a quintessential aspect of the ECI, but also is the most troublesome when adapting the technique for CMC. Three specific topics will be examined in subsequent analyses including the principle of synchrony, the nature of interruption, and the use of continuers.

The principle of synchrony (Matarazzo and Wiens, 1985) is related to the phenomenon of emotional contagion whereby collocutors mirror one-another's behaviour over time. The interviewer, wishing to maintain a calm and casual rapport with an interviewee may then regulate their nonverbal behaviour in order to 'transfer' the desired affective state. An important research question for this study is how could the benefits of such tactics be reproduced through CMC, where many of the conventional nonverbal behaviours are masked or transformed?

Interruption is an easy trap to fall into, as was illustrated by Annette Markham's ethnography. The interviewer can become anxious waiting for the interviewee to respond or can change the topic prematurely. Anticipation of the transition-relevance place, strengthened by the interview structure per se, is also weakened by the concealment of nonverbal signals. Additionally, nonverbal continuers may need to be replaced by verbal ones, which itself can cause an interruption, possibly promoting a conversational norm of interruption. Conversely however, it is impossible for the interviewer to interrupt the interviewee mid-utterance, as all the parts of the utterance are posted at the same time-index and cannot be subsequently broken up. An important question is how can the interviewer avoid the disruption of rapport through the



reduction of interruptions, whilst still providing the increased number of verbal continuers that are needed in order to 'listen through words'?

It is also important to explore how the interviewer can tell when the interviewee has completed what they were going to say, and what continuers and other similar conversation devices provide the impression that more information is sought, rather than the impression that a subsequent utterance is about to be delivered (i.e. to avoid a situation where both speakers are waiting for the other to respond).

The following guidelines are suggested:

- *Demonstrate patience and wait for a response before prompting.*
- *Do not use newsmarks that are ambiguous without vocal paralinguistic markers, unless they are coupled with punctuation that provides insight to the intended speech act.*
- *If two utterances are posted simultaneously then they may act in competition. In such a competition, the interviewer must try to explicitly withdraw (although this might prove difficult if the interviewee has started to consider the new utterance).*

### **7.3.2 Phase 2: Explaining the Interview Aims**

This phase is conducted face-to-face before the rapport building stage of phase one. The need to build rapport whilst online dictates that this should be the case. As with a standard ECI, the interviewee will be asked to report everything they can remember and the interviewer will need to promote focused retrieval. Additionally, the interviewer must stress that the interviewee should be in control of the interview, and that the interviewer is merely a facilitator in the role of eliciting memory (Fisher and Geiselman, 1992). These points will be reiterated at the beginning of both the free narrative and the questioning phases.

### **7.3.3 Phase 3: Initiating a Free Narrative**

Context Reinstatement involves a momentary detachment of the interviewee from the interviewer. This detachment could result in a greater feeling of isolation through CMC than it would face-to-face. In effect, this detachment means that the interviewee has to disconnect themselves entirely from their sole means of communication with the interviewer. This disconnection represents a momentary change of context, and this transition needs to be managed carefully so as to minimise the disruption to rapport. This management becomes difficult if the communicative detachment between the two co-speakers has enlarged. Further investigation needs to be carried out into the verbal techniques that could be employed to provide a smooth transition both into and out of mental reinstatement without any detrimental effects to rapport.

If the interviewee does not receive sufficient information of how long they are to spend mentally reinstating the problem, then they might speed up the process in order to re-establish contact with the interviewer. Alternatively, monitoring the CMC software to ensure that no subsequent utterances have been posted may distract them. The extent to which this monitoring actually occurs is a subject for further investigation and it is recommended that the interviewer suggest a suitable period of time to spend on the reconstruction activity.

After the interviewee has spent some time reinstating the context, it is then necessary for the interviewer to attempt to elicit a lengthy free narrative. For the CMC interviewer to obtain this without interruption is particularly difficult. The free form, uninhibited, conversational state of the interviewee needs to be maintained, and so it is not desirable for them to simply word-process a statement into the computer. Within any QS-CMC interaction there will always be the potential for some degree of editing that will be masked from the interviewer. However, in a full-length narrative this process of revision could restrict the naturalistic flow of recollection. Typing speed may also impede this flow, and so wherever possible it is important for the interviewer to promote an uninhibited flow of discourse.

During a face-to-face narrative, once an utterance has been spoken, it cannot then be changed. In CMC an utterance can be changed right up to the point at which it is posted and it is fair to assume that the longer the utterance, the more modifications may occur. It is possible that as the account grows, the interviewee may become aware of the unstructured sequence of their earlier statements and this could result in excessive editing. Such a freedom from imposed form is necessary for CBCA to be successful and is therefore an important norm to promote within the phase. Encouraging the interviewee to post more often can reduce the average utterance length. However this would also increase the number of turns and consequently the number of transition-relevance places. This form of interaction then starts to look similar to typical CMC conversation, and the difficulty may be that the interviewee slips into this discursive style. In such cases the interviewee may not continue their utterance until some response (such as a continuer) is received from the interviewer. However it is difficult to judge when such a response is desired, or when the interviewee is simply thinking or typing. The overall problem is that the interviewer may get drawn into the narrative, resulting in an undesired level of interference.

### **7.3.4 Phase 4: Questioning**

Many of the questioning principles employed within the ECI can be applied through CMC. The guidelines concerning open and closed questions and regarding activating and probing a mental image remain. However, the persistence of the uttered question in CMC gives it a permanence that is of a very different nature than the spoken utterance. This persistency provides the interviewee with an up-to-date history of all they have said and this can have a number of implications for the method per se, as well as its inclusion as a key part of SVA. For example, an accessible history might reduce inconsistencies for a deceiver and thus improve their CBCA

score. Alternatively, the very persistence of a written record may reinforce and strengthen the perception of an only vaguely remembered but incorrect piece of information.

When an interviewer poses a question, it is important for that interviewer to assess at that moment whether or not the question has been understood. Cues to the perception of misunderstanding may be verbal or nonverbal. Upon the perception of such cues, the interviewer may reformulate, exemplify or take some other form of action, such that the misunderstanding is quickly resolved and rapport is maintained.

Within CMC, nonverbal signals of miscomprehension are all but lost. Verbal signals come in two forms, the explicit response to restate the question and an answer to a misperceived question. The latter form of response is where difficulties can arise. The difference is again due to the way in which QS-CMC utterances are formulated whereby the listener does not have access to the construction of the utterance. This means that the miscomprehension goes undetected until the entire response has been formulated. The problem for the interviewer is how to respond to this utterance. Simply ignoring it and asking the interviewee to reply differently may cause embarrassment. Additionally, if the interviewee has spent time thoughtfully constructing the response then it should not be disregarded out of hand. This could indicate that the interviewer is not interested in the new information or possibly that the participant has not provided the 'correct' answer that was expected by the interviewer. Such actions may increase the participant's vigilance over their responses, damaging the free flow of discourse.

Within QS-CMC, the interviewee has access to an accurate transcript, composed in real-time and displayed in a history frame. This transcript is useful for the interviewee to review, and see if anything has been omitted from their account. However from the perspective of CBCA, it will aid in the memory of what has already been said, thus reducing the relevance of the '*Logical Consistency*' criterion. The ways in which the message history may be used to help or hinder the ECI is an important topic to be explored within this phase.

### **7.3.5 Phase 5: Varied and Extensive Retrieval**

It seems possible that the interviewer can use the same types of varied retrieval strategies in CMC as they would in a face-to-face setting. Two strategies of varied retrieval will be attempted in this study. Interviewees will be asked to report details in a reverse order of time and to report details from another's perspective.

### **7.3.6 Phase 6: Summary**

The summary stage of the ECI is a peculiar one to be conducted over CMC. The message history already provides the interviewee with a full record of the interview. However this is not a *summarised* account and it may still be useful for the interviewer to provide a recap. In the CMC environment, this could be fashioned as a single large utterance by the interviewer, posted at a single time interval for inspection by the interviewee. This could cause a number of difficulties

as well as resulting in a detachment with the interviewer for a prolonged period. For example, the interviewee may feel pressured to read the summary as quickly as possible and this becomes a problem if the interviewee needs to critique the whole of the summation at the end, as opposed to interjecting intermittently. They simply might forget their comments concerning the earlier parts of the summation, or may focus solely on one particular error or image.

The summation might also be achieved by asking the interviewee to review the message history. It remains to be seen how effective this task will be, as once again this draws the interviewee away from the interviewer and there will be a pressure for the interviewee to reconnect. This might also be quite a complex task, considering that the interview is neither edited nor summarised. It remains to be investigated how the summation can be presented to the interviewee in such a way as to promote the addition and clarification of information. The solution to this problem could be technical, perhaps through the highlighting of relevant passages in the interviewees chat history frame.

### **7.3.7 Phase 7: Closure**

Interview closure will occur both offline and online and this stage of the ECI blends in with the debriefing phase of the research protocol. A smooth transition from the online environment, back to the face-to-face environment is necessary in order to maintain the fluidity of the interaction. The interviewer will need to use termination sequences that are very definite. This will be helped by the software, which will inform the participant, after the termination sequence, that the interviewer has disconnected from the system.

### **7.3.8 A Summary of the Explorative Themes**

Eleven key themes for exploration were outlined in this section. These themes are considered in the following analysis and are summarised as follows:

1. *How effective is the Free-Narrative Phase?*
2. *Can paralinguistic techniques be employed within CMC that convey empathy and listening?*
3. *Can the interviewer avoid interruption whilst still providing a sufficient number of prompts such as continuers and newsmarks?*
4. *Are the aims of the interview and phase instructions expressed effectively?*
5. *How can the transition in and out of mental reinstatement be managed?*
6. *How much time is spent reinstating the context?*
7. *How effective is the questioning process?*
8. *How are misunderstandings resolved?*
9. *How is the message history used?*
10. *How effective are the varied retrieval methods?*
11. *How effective is the Summation Phase?*

## 7.4 Method

The adapted ECI was examined by conducting a theory-driven discourse analysis on the interview transcripts. Nine participants were selected as an opportunity sample of experienced computer users that were familiar with CMC. Initially, participants were greeted and provided with the participant information sheet (see Appendix III) before signing a consent form for the study. The protocol was comprised of three stages that involved the participant watching a videotape of a short sequence<sup>1</sup>, and then being tested for knowledge of that sequence through a CMC based ECI. Upon conclusion of the ECI, a short unstructured interview was carried out with the participants to explore their experience of the process.

The author wrote a simple QS-CMC chat-system in order to conduct the ECI. The system was provided with a facility that recorded the entire transcript of the conversation along with timestamps for when each utterance was posted. A screenshot of the interface to this system has been provided in the Appendix. The system comprised of a server process that managed the promulgation of messages from distributed client processes that adhered to a standard protocol. Only one, basic type of client was developed for this study and was used by both the interviewer and interviewee alike.

## 7.5 Analysis

The transcripts were analysed in discourse-analytic and qualitative manner. The transcripts were analysed phase-by-phase examining the general operation of the ECI and by examining the explorative topics stated in section 7.3.8.

### 7.5.1 Free Narrative

At the beginning of the free-narrative phase, the interviewee was asked to spend a few minutes visualising the video clip in their mind. They were told to signal the interviewer once a sufficient period of time had elapsed. It was hoped that this would allow the interviewee to spend the appropriate amount of time reinstating the context without feeling pressured to finish too quickly or feeling the need to pay attention to any utterances posted by the interviewer. The chat-software recorded the time spent on this activity. On average, the interviewees spent about 34 seconds (SD 22) between receiving the request and signaling that they had finished thinking about the clip. It would seem then, that the participants did not spend the necessary time required in order to fully reinstate the context. The reason for this could lie in the unease caused by that momentary disconnection of the interviewee from their sole means of contact with the interviewer.

---

<sup>1</sup> This sequence was taken from the motion picture 'The Spanish Prisoner' (Sweetland Films, 1998).

Following the context reinstatement, the participants were asked to relate the details from the video clip from start to finish. Typically, the participant would break this narrative into a series of postings as is shown in the following example.

56	E/W	Then the next scene is in a CCTV room
57	INT	The younger exec is asking about CCTV procedure, and is told all tapes are stored in case of insurance idsputes
58	E/W	Distputes
59	INT	good so far
60	E/W	Then he is coming back from town with a new suit in a car, and gets out and talks to a woman
61	E/W	Who is at least superficially very bubbly and friendly to him
62	E/W	They meet again while he is walking along the shore for a second time
63	E/W	A flying boat lands in the background
64	INT	what else?
65	E/W	She says something about the Caribbean being the main tourists destination of Japanese honeymooners
66	E/W	"Strage considering the distant" says bloke
67	E/W	"Must be cos it's so romantic" flirts the bird
68	INT	:)

This example illustrates the three main types of continuers used by the interviewer within the free-narrative sections, appraisals (line 59), information requests (line 64), and smileys (line 68). The post-hoc interviews indicated that by using verbal continuers, the interviewer did manage to enhance the perception of 'listening'. Nonetheless, appraisals can also enhance the perceived accuracy that an interviewee attributes to a memory and so should always be used with caution. Similarly, information-requests may act to rush the interviewee causing them to skip-over important details. Smileys can also express an emotional reaction to the information recollected and can have the same effect as appraisals. Whilst the interviewees indicated that they had not evaluated these continuers in any depth, other more neutral forms of continuation mechanism may be more suitable.

One participant departed from the multiple-post method and provided their entire narrative within a single posting. A post-hoc interview with the participant (who did not have any experience with chat room software) revealed that they had found the process to be rather un-interactive, likening it to a rigid 'question and answer' session. They also reported the temptation to 'skip-over' details and to mention only the main points from each scene. It might be the case that the multiple-post method has distinct advantages over the single-post strategy. These could include:

1. *Breaking up the narrative to discourage extensive editing and enhance spontaneity.*

2. *Reducing tedium through the enhancement of interactivity.*
3. *Breaking up the text to lessen the perceived effort of writing the narrative, splitting one big task into many small tasks.*

Interviewees would typically end their narrative with some sort of exclamation *"that's all I can remember"* or a more definite *"the end"*. The second exclamation reflected the story-like scenario that was used as the stimulus material. At the end of the narrative, the interviewer asked whether or not there was anything else that the interviewee could remember. Usually, no more information was presented until subsequent phases. The second phase began with a short appraisal, *"okay that was good"* before instructions for the next phase were provided (all in the same post). One interviewee mentioned that this appraisal seemed like it was "cut-and-pasted-in" and this was a correct observation. The appraisal was placed into the same utterance as the questioning phase instructions in order to self-select and bridge the two phases together. It may have been better to break this utterance into two parts, allowing the interviewee to respond in-between. So for example, a short question asking the interviewee how they felt could follow the utterance *"okay that was good"*. The instructions for the next phase could then be sent in the following post. This might then improve the sense of interactivity by making the situation less artificial, whilst still maintaining control of phase transition and clearly indicating the direction of the interview.

Six of the interviewees used the present tense when describing their accounts (this can be seen in the example above). Use of the present tense might indicate that the interviewee is deeply involved in the recollection process, describing what they are seeing in their mind as they see it. Alternatively, the use of the present tense might in some way be related to the stimulus type, a videotape as opposed to a real-life experience.

Post-hoc interviews revealed that despite finding the experience to be a bit strange, the interviewees did eventually become accustomed to the narration process. One interviewee reported becoming quite engrossed in the construction of the narrative. Another participant mentioned having too many *"balls to juggle"* because they needed to concentrate and type at the same time. Typing and reading is typically a more conscious process than speaking and hearing and so there is a danger of cognitive resources being diverted away from remembering and into the mechanisms of utterance construction and interpretation instead. It might be surmised that this would only become a problem if the interviewee felt pressured to respond and it was found that this was the case for some.

It seems as though the low level of interactivity and the nature of feedback is an important consideration during the free-narrative stage. Although comprehensive free-narratives were obtained in each of the interviews, it seems as though this process might be helped if more perceptual affordances were provided within the chat room. Such affordances may help to assure the interviewee that someone was listening to them and could provide feedback as to the appropriate level of detail to relate.



7.5.2 Questioning

The transcripts indicate that the interviewees understood the instructions given to them at the beginning of the questioning phase. The interaction within the questioning phase typically followed a question-answer format, and only rarely did the interviewee interrupt with new information that did not pertain to the immediate question. A standard question-answer sequence is easier to manage within CMC than a more colloquial sequence that can include interruptions. This latter type of sequence is promoted within face-to-face interviews where the interviewee is not forced to speak about specific topics at certain times in such a way that any spontaneously remembered details are then forgotten. However the informal sequence mentioned above is not as easy to manage through CMC as is the question-answer sequence, and therefore cognitive resources are taken away from the mechanics of memory retrieval and diverted into the process of conversational management and repair. Interruptions for example, can cause more disruption within CMC because the repair processes are starved of the real-time feedback that they need in order to make informed behavioural adaptations. These processes are either slowed, or can be further confounded if the pressure to respond is great, and subsequently repair behaviour is rushed with choices based upon insufficient feedback.

Questions were successful in obtaining a moderate amount of new information, especially relating to the appearance of characters or the details of locations. The following example illustrates a questioning sequence concerning one of the characters in the clip.

97	INT	okay then, I would now like you to describe the women on the beach, what was she wearing, what did she look like?
98	E/W	she was wearing a hat 1920's style har
99	E/W	red floral outfit i think
100	INT	1920's style?
101	E/W	you see women in old black and white photos wearing them -> for example the queen mum
102	INT	okay, what else do you remember about her appearance?
103	E/W	she was shorter than him dark hair
104	INT	anything else? What sort of things did she sayt to him?
105	E/W	well there was the quizzing about the meeting he just had and she seemed very interested in him
106	INT	Why do you say that?
107	E/W	well she made a comment about how romantic the caribbean was and when he didn't even smile she looked dissapointed
108	INT	Good, is there anything more?
109	E/W	hmmm
110	E/W	no it's slipped
111	INT	that's okay



In this example, the questioning brought out new information about the woman's hat, hair colour, and height. Further probing revealed additional features of the spoken exchange between the women and the man on the beach (lines 105, 107). The questioning sequence for this image ended in a manner that is similar to that found in spoken interviews. After a period of about forty-five seconds, the interviewer posts a final probe for more information (line 108). After a thinking period of about twenty-three seconds, the interviewee seems to 'buy more time' by posting their utterance on line 109. The expression 'hmmm' performs the act of informing the interviewer that the interviewee is both aware of the question and needs more time to consider a response. The final response is then given fifteen seconds later on line 110. The presence of the utterance on line 109 functions to maintain contact with the interviewer and to promote interactivity. This is performed through the verbal expression of a commonly known vocal behaviour for 'pondering'. Another interviewee used a similar turn-reservation device during the free narrative:

87	INT	yep, I understand
88	E/W	Anyhow, hmm...
89	INT	anything else?
90	E/W	She says something, no idea about what she says

Some interviewees heightened the sense of interactivity by answering questions on multiple lines. Where lots of information was placed within a single utterance, it was often written in a very tight form without many connecting words or phrases such as in the following example.

57	INT	You mentioned the younger man, who featured throughout the clip. Try and picture him in your mind. I would like you to describe him for me in terms of his appearance, his manner etc. Try to give me the most detailed description that you can.
58	E/W	Medium height, slimish build. Brown hair, slightly curly, often ruffled by the sea breeze when outdoors. White or very pale cream shirt, with (darker not tan) brown trousers and jacket. Darkish tie, presumably matching suit in some way. Often carried his jacket over one shoulder. I think maybe wore glasses.

Confabulations occurred infrequently throughout the questioning phases. In the example above, the younger man's trousers were grey, not brown and the tie had red stripes and was not particularly dark. Similar errors were made regarding colour throughout the clip. Two of the interviewees recalled a boat at the end of the clip as being coloured white when it was orange. The way in which one of the interviewees expressed this is noteworthy because it gives an indication of where the misperception arose.

145	INT	Can you describe the boat?
146	E/W	small power boat - white, I think (everything seems to be white!)

7.5.3 Varied Retrieval

The two methods of varied retrieval used within the interviews involved asking for details about the sequence in reverse order and asking the interviewee to tell the story from the perspective of a character within the clip. These varied retrieval questions were not particularly successful in generating new information but they did not result in greater confabulation either. In particular when participants retold a sequence in reverse order, they mirrored what had been written before in the free narrative as in the following example.

Free Narrative			Varied Retrieval		
61	E/W	Then she makes a general comment on the caribbean, this confirms the location of the film.	157	E/W	The brunette walks back to the hotel
62	E/W	she states that its the number one holiday location for young japanese.	158	E/W	Jo takes a picture of the brunette
63	E/W	he states that its a bit far.	159	E/W	the picture is of here in front of the beach and the sea, quite possibly with the moured boat and the man in the background
64	E/W	Moving towards the beach	160	E/W	The brunette takes 2 photo's of Jo by the Sea
65	E/W	oh and a boat (more like a dingy) being moured.	161	INT	good...please continue like that
66	E/W	she takes her camera and takes two photo's of Jo	162	E/W	Jo and the Brunette talk on the beach
67	E/W	he then takes his camera out	163	INT	talk about what?
68	E/W	offering to take her picture, and finally doing so	164	E/W	they mention the caribbean and how it is the most visited holiday destination for young japanese
69	INT	good so far	165	E/W	Jo mentions how this is really far
70	E/W	meanwhile the people in the boat seem to be disembarking			
71	E/W	she moved away back to the typically colonial building (hotel)			

This could be indicative of the interviewee simply scrolling back to see what they had previously said and retelling the same information. This is consistent with reports in post-hoc interviews that revealed that the interviewees were becoming tired at this stage (in the final two interviews,



the varied retrieval questions were intermingled within the questioning phase and this seemed to lessen the interviewees aversion to the task). This style of retrieval did generate missing dialogue with two of the interviewees and resulted in the telling of a 'missing scene' by another.

7.5.4 Summation and Closure

The participants were asked to review the message history upon conclusion of the interview and to fill in or correct information. On average they spent about 49 seconds (SD 24) on this activity, a similar length of time to the time spent during context reinstatement. Four of the participants did not add or correct anything, but the other five did have new information to add. Given the length of the transcripts between the start of the free narrative and the point of closure, it seems unlikely that the interviewees did not actually read through the transcript. At this stage of the interview, the interviewees may have been relatively tired and might not have invested the effort required to thoroughly review the transcript. This effort could be reduced by a number of methods. A function could be implemented within the chat system that allows the interviewee to view only the statements that they themselves have authored. A highlighting function could be used so that the interviewer could highlight certain areas of the transcript revealing the pertinent information in the transcript and occluding less relevant information. Alternatively, the interviewer could simply cut-and-paste a summary of the facts into a window and that information could be shared with the interviewee during the summation phase or during the interview (as a shared white board).

Post-hoc interviews also revealed that the interviewees perceived that the session had passed much quicker than it actually did. This is some indication that the interviewees were relatively engrossed in the session.

7.5.5 Misalignment and Repair

There were five instances of conversation misalignment throughout all of the transcripts and this was much lower than was expected. It is possible that the orientation to the question-answer structure was responsible for this. Misalignment, when it did occur, was usually resolved fairly rapidly. The most disjointed instance of misalignment is provided below.

18	INT	Now we are going to begin to talk about the clip.	10.99
19	INT		15.21
20	E/W	ok	2.64
21	INT	Sorry, last the second half of that (may have sounded a bit brash!)	26.91
22	INT	lost rather	6.04
23	INT	What I'd like you to remember throughout the rest of the interview is that *you're* in control! At any point if you remember something "out of the blue" then just cut in	11.54



		and tell me!	
24	E/W	no it didn't	2.47
25	E/W	right - will do	10.38
26	INT	good, thanks!	0
27	INT	What I would like you to do, is to try to visualise the video clip that you've just seen. Try and recreate the clip in your mind, scene by scene. Take a few minutes to do this and send me a message when you've had time to reflect.	10.38
28	E/W	right	33.94

In line 18, the interviewer has accidentally posted their utterance before it was completed; in line 19 the blank response was another mistake. The interviewee has interpreted a TRP at this point, replying with a newsmark. In line 21, the interviewer apologies for the error and subsequently makes another spelling mistake (indicative of the heightened pressure to ensure that the interviewee was not offended by a seemingly blunt response). The interviewee responds to the utterance in line 21 by line 24 by which the interviewer has already provided a self-repair (line 22) and the second part of the utterance begun in line 18 (line 23). A phantom adjacency pair occurs in lines 25 and 26 but this does not cause any difficulty and by line 27 the misalignment has been corrected. This sequence demonstrates that both the interviewer *and* the interviewee are trying to orient to one another. The norms of adjacency mean that the interviewee's performance in line 24 has the potential interpretation as a response to line 23, and subsequently a quick separate confirmation is provided that functions to both to acknowledge the utterance in line 23 by means of a newsmark, and also to re-contextualise the utterance in line 22. This performance assumes that the interviewer will be able to deduce from the timing of the postings and their content which response matches up with which statement.

The interviewer's utterance in line 22 illustrates the principal type of repair employed within the interviews whereby a misspelling is clarified by simply repeating the corrected word, sometimes with a short word or phrase that performs the function of making the error explicit. The interviewees made nine instances of this type of repair and some of these are repeated below.

44	E/W	Meanwhile a flying bot comes over the island and into land while they watch. As the man and woman? come ashore from the boat Joe and woman are standing on beach.
45	E/W	flying boat even :)

138	E/W	Hmm... not sure. Just got that impression. Like he just appeared out of nowhere... perhaps it was the way the director cut straight to his face, without including his approach to the younger guy.
139	E/W	sly - stealthy.

57	E/W	The younger exec is asking about CCTV procedure, and is told all tapes are stored in case of insurance idsputes
58	E/W	Distputes

It seems as though conversational misalignment and repair, thought to be highly haphazard within some CMC contexts, might not be problematic within a dyadic investigative interview.

## 7.6 Summary and Discussion

This section is comprised of three parts. Firstly, the main findings of the study have been summarised in section 7.6.1 with respect to the explorative topics presented earlier. The discussion then progresses to an evaluation of whether the adapted ECI provides a sufficient framework to be used as part of a CMC-based SVA in either its current or improved form. The discussion is concluded by suggesting several improvements to the interviewing technique. These changes will then be incorporated in order to ensure that the technique can be used as a part of the Statement Validity Procedure detailed in the next study.

### 7.6.1 A Summary of the Findings by Theme

#### *7.6.1.1 How effective is the Free-Narrative Phase?*

The free-narrative phases were successful for each interviewee. The interviewees found this style of discourse unusual at first because it did not resemble the way in which CMC is normally used. Narratives were more detailed when posted on multiple lines and it would seem that the interviewees should be encouraged to do this. Some interviewees used the present tense when providing their descriptions and this could be indicative of concentration. Interviewees were able to successfully indicate when they had finished their descriptions. Some interviewees expressed the difficulty of both remembering and typing at the same time.

#### *7.6.1.2 Paralanguage, Empathy and Listening*

A sufficient number of verbal prompts were employed to enhance the perception of listening although some interviewees still felt that the interviewer was not listening. Paraphrasing the interviewee during the questioning phase was effective, but sometimes simply paraphrasing with a question mark makes it unclear whether the interviewee is seeking further clarification, or does not understand the utterance. Verbal newsmarks and continuers can easily be construed as evaluative and should be prepared beforehand. Some, less evaluative newsmarks, if not 'banal' may not be understood. This type of misunderstanding might be reduced if the interviewer describes some of the newsmarks that will be used before the interview begins and explains that this is the convention that will be used to indicate the receipt of an utterance.

### **7.6.1.3 Avoiding Interruptions**

Interruptions rarely occurred. When misalignment did occur it was quickly dealt with.

### **7.6.1.4 Expressing the Interview Aims and Phase Instructions**

The interviewees did not find any difficulty with the instructions. Sometimes further clarification as to how much detail to provide in the free-narrative stage was needed. Initially, varied retrieval instructions pertaining to assuming a character's perspective were misunderstood. These were clarified and subsequent interviewees understood these modified instructions. Cut-and-pasted instructions were sometimes perceived as such and this damaged the perception of listening. There is an advantage of cutting and pasting questions and perhaps this process should be made clear to the interviewee.

### **7.6.1.5 The Transition into Mental Reinstatement**

The transition into and out of mental reinstatement was effective. Some interviewees required further clarification as to how they were to signal to the interviewer once they were ready.

### **7.6.1.6 How Much Time is Spent Reinstating the Context?**

Insufficient time was spent reinstating the context. It might not be suitable to conduct this portion of the phase online. The adapted ECI might work best when offline phases and online phases are combined. However, the interviewee is likely to feel less embarrassed, closing their eyes and visualising the context when they are alone, than when they are in front of the interviewer, and so perhaps the interviewee should be given a fixed time to reinstate the context. The interviewee may then need to do some 'clock-watching' but this still may promote a longer period of context reinstatement.

### **7.6.1.7 How Effective is the Questioning Phase?**

Questioning was effective and answers were usually posted as a single utterance or using multiple utterances. It is possible that the second method is better, resulting in a good balance between reducing over-editing whilst still promoting uninhibited recall. The questioning resulted in only a minor number of confabulations and these were predominantly centered upon the colour of objects. Questioning was typified by an orientation to 'question-answer' sequence. This structure has the potential to facilitate rather than impede recall by reducing the effort required to locally manage conversation.

### **7.6.1.8 Resolution of Misunderstanding**

Misunderstandings were infrequent and repair processes were effective. Where such misunderstandings did occur, there was evidence to suggest that repair processes could be performed by both the interviewee and the interviewer. Interviewees would often use verbal

signals to indicate they were thinking, and this acted to prevent misunderstandings with the interviewer.

#### ***7.6.1.9 How is the Message History Used?***

There was some indication that the message history was potentially hindering the effectiveness of the varied retrieval phases. It would also seem that the history was not being sufficiently perused upon conclusion of the interview.

#### ***7.6.1.10 How Successful were the Varied Retrieval Methods?***

Varied retrieval methods were largely ineffective, but importantly did not lead to confabulation. It could be surmised that these methods are less effective in scenarios where the interviewee is not fully immersed in the event.

#### ***7.6.1.11 How Effective is the Summation Phase?***

The summation stage was effective in some cases even though it would seem that interviewees did not look over the entire message history.

### **7.6.2 Is the Adapted ECI Suitable for SVA?**

The version of the ECI developed within this study seems appropriate to be used with SVA. This is primarily because all of the interviewees yielded detailed narratives, with little confabulation and responded to effective, non-leading questioning. In addition, fears of potential misalignment and misunderstanding would seem to be unwarranted. The CMC-based technique has the advantage of reducing the negative effects of nonverbal behaviour whilst effectively compensating the drawbacks through the orientation to semi-structured sequences. In addition, the careful use of technological aids helped the interviewer to schedule and to overcome the difficulties of utterance construction.

This study was a quasi-experimental study that used a small sample size and only a single interviewer. In this regard, the study is limited to the finding that such interviews are possible, for a particular interviewer and for a particular set of interviewees (predominantly postgraduate, computer-literate participants). However, this qualitative study has succeeded in its exploratory aims by describing some of the performances that occur within the CMC-based ECI procedure. Such information is valuable to provide both an interviewing framework, and a selection of factors to explore within subsequent large-scale studies. These studies could examine how the interview might be automated further (or even entirely) and how interviewees might respond in different ways to the process.

This study has revealed a number of issues that will need to be altered before the technique is used in the SVA procedure in the next study. In addition, the experience of interviewing has revealed some ways in which the technology might be adapted to facilitate the process. These changes are discussed next.

### 7.6.3 Improvements to the CMC-based ECI

The purpose of this study was to theoretically derive a CMC-based ECI, and to empirically investigate ways in which this technique could be improved. This would ensure that an ECI interview would generate transcripts of a sufficient quality to be used in the SVA procedure reported in Chapter 8. The main findings that have been found from the discourse analysis have been outlined in the previous sections. Within this section, these findings are used to describe certain key changes that will be made to the ECI procedure in preparation for the next study. These changes have been made in response to the findings examined under the themes in section 7.6.1 and in response to other observations by the investigator.

Themes 1 and 2 revealed that the interviewee is unsure as to whether or not they are being listened to, especially during the free narrative phase. It is suggested that:

**At the beginning of the interview, it should be explained to the interviewee that it is quite normal to feel that they are not being listened to, and that this is due to the lack of nonverbal feedback. They also should be told that despite this feeling, the interviewer is highly focused upon their narrative and responses but sometimes make take a while to respond in order to spend the appropriate amount of time reflecting upon them.**

Theme 4 revealed that interviewees are aware of pre-fabricated question stems that appear to be cut-and-pasted into the chat entry box. This technique helps the interviewer but could damage rapport by displaying poor active listening skills. It is suggested that:

**At the beginning of the interview, it should be explained to the interviewee that question stems will be cut-and-pasted into the chat entry system. It should be further explained to the interviewee that the purpose of this is to minimise the waiting time for the interviewee, and that not all of the questions will be copied in this fashion.**

Theme 2 revealed some difficulties in comprehending newsmarks and continuers. It is suggested that:

**The interviewer should not use the paraphrasing technique that simply re-quotes with a question-mark suffix as this can cause confusion. In general the interviewer should try to use non-evaluative newsmarks such as “okay” and “I see”. However, when the interviewee tells a joke or mentions something adverse, the appropriate level of empathy should be conveyed using simple comments or emoticons.**

Themes five, six, and eleven revealed further ways in which the context reinstatement and summation might be accomplished.



**The interviewer must provide the interviewee with a fixed time period in which to reinstate the context and to review the summary. The interviewer should re-establish contact with the interviewee once the required period has elapsed.**

Themes four, nine, and ten suggest the limited value of certain varied retrieval phases within the CMC-based ECI. It is suggested that:

**The interviewer does not asked the interviewee to recount their story in reverse order of events as this narrative can be corrupted by the message history. The instructions used to instruct the interviewee to assume the perspective of another character should be clarified by adding the instruction “telling me the story as if you were X”.**

Finally, it is suggested that pressure should be taken off of the interviewer by providing them with additional software. This software should provide facilities for note taking, utilising question stems, and for keeping time of each phase of the interview.

All of these changes will be made to the CMC-based ECI in order to increase its effectiveness and appropriateness to be used as part of a Statement Validity Assessment. This study has shown that this technique provides an effective first component of the SVA truth-validation procedure. Subsequently, the improved version of the ECI will be used in the study described in the next chapter.

## Chapter 8

# A Statement Validity Assessment within CMC

A Statement Validity Assessment was carried out using a CMC-based interviewing technique. Twenty-five participants were asked to report one truthful and one fabricated story of a day trip. The transcripts were coded by two blind raters using the Criteria-based Content Analysis (CBCA) technique. In addition, the CMC software recorded information regarding what the interviewee had deleted from the chat-entry box. It was found that truthful statements were correctly identified by one rater 76% of the time and by a second rater 80% of the time (using a subjective total measure). However, a more thorough analysis of criteria frequency showed that CBCA demonstrated no significant univariate or multivariate discriminatory power in evaluating the truthful statements from the fabricated statements. Many of the participants managed to use countermeasures that attacked some of the specific criteria used by CBCA. In addition, no significant differences were found in the deletion behaviours of truth-tellers and fabricators. It is suggested that CBCA in its present incarnation is ineffective at discriminating against truthful and fabricated statements generated through QS-CMC. Considering the moderate support CBCA has received in the face-to-face domain, this chapter is concluded by suggesting some possible explanations for the refutation of the hypotheses within this context.

---

## 8.1 Background

Effective nonverbal skills would seem an indispensable part of the investigative interviewer's repertoire. Through the careful regulation of nonverbal behaviours, the interviewer can maintain good rapport with the interviewee, control their affective state, and promote recall using active listening techniques (Milne and Bull, 1999). By attending to nonverbal signals, an interviewer can regulate their behaviour to convey empathy, and monitor the interviewee's affective state in order to provide them with reassurance. Nonverbal processes are central to the regulatory, reciprocal, and adaptive functions of everyday communication and therefore their presence within an investigative interview would seem essential. However, the research into mediated communication shows that the conveyance of nonverbal cues is not a necessary mechanism for the accomplishment of good rapport, active listening, and successful interaction. This is an important finding because it has the potential to unshackle the SVA procedure from the face-to-face domain, allowing the interviewer to leverage the opportunities presented by mediated forms of communication.

The extent to which interviewers differ in their ability to obtain an eyewitness statement is a central concern for SVA research and practice. Any qualitative differences between transcripts that have systematically arisen from a particular interviewer's style have the potential to confound CBCA or RM. There are many such dangers that can transpire from the

misapplication of nonverbal skills. The stress on the interviewee can be heightened resulting from a poor application of proxemics, the interviewee responses can be corrupted by the semaphoring of appraisal, and rapport can be damaged through a poor understanding of cultural differences in turn-taking and the misinterpretation of paralanguage (Milne and Bull 1999). SVA research has attempted to reduce the potential for such variability by devoting effort to standardised training programs. However, the use of nonverbal behaviours is not an exact science, and it is difficult for training programs to completely eliminate the diversity that exists between the basic expressiveness and social skills of different interviewers (factors that heavily influence impression formation; Riggio and Friedman, 1986; Friedman et al., 1988). In addition, the physical presence of the interviewer can make the interviewee more conscious of their nonverbal self-presentation, especially if they are suspected of committing a crime. Witnesses who are particularly shy or withdrawn may find it difficult to confront an interviewer face-to-face and this would be detrimental to effective recall.

It is suggested here that whilst the appropriate training of SVA interviewers is an essential prerequisite that promotes standardisation, in some circumstances the presence of nonverbal cues and the presence of the interviewer can be detrimental to the SVA process. By creating a systematic variability of the quality of transcripts attained by different interviewers, studies that evaluate CBCA and RM can become difficult to compare. Moreover, it is suggested that the removal of these cues will make interviewer training simpler and will increase standardisation. Nonverbal cues are not essential for good rapport, and these can be compensated for within CMC by all but the most egregious interviewer. In addition, a medium such a CMC has a number of advantages over face-to-face communication. Specifically, the removal of nonverbal cues and copresence can make a withdrawn or shy interviewee more forthcoming by removing the emphasis on difficult physical self-presentation, and by focusing the interviewee on the contents of their narrative and responses. It is speculated that difficulties resulting from cultural differences in turn taking and paralanguage will be quashed due to the superordinate conversational regulation processes of CMC. It is suggested that a CMC-based SVA might be a useful adaptation of the technique for co-operative and computer literate witnesses as well as being a useful alternative to be used with shy or withdrawn witnesses.

The investigation of a CMC-based SVA also has another purpose. Little is known about the content of deceptive encounters within CMC. Whilst it may be fair to assume that qualitative differences exist between the contents of face-to-face and CMC exchanges, it is interesting to investigate whether this confuses the differences posited by CBCA and RM. Whilst these coding strategies are not appropriate for naturalistic face-to-face lie-detection, this might not necessarily be the case with naturalistic CMC lie-detection. This study is a useful step towards such an ambitious goal because under some circumstances naturalistic behaviour within CMC can be far more stereotypical than it is face-to-face. Even if this stretches the applicability of content analysis too far by breaching the constraints of the validity checklist and delving too far into positivism, it is still important to examine what the effect of such a transposition may have on the CBCA and RM criteria.

This study is concerned with the transposition of SVA using transcripts generated from CMC interviews. This transposition will be discussed further in the remainder of this section before the hypotheses posed by the study are stated. The adaptation of the SVA interview was the subject of the previous chapter and these findings will be summarised next. Then the content criteria will be re-examined with regard to how they might be affected by the CMC medium. This section will conclude by hypothesising what nonstrategic leakage cues might be used to complement the content criteria.

### **8.1.1 The Interview**

The technique of CMC investigative interviewing was explored in the previous chapter. Some pertinent critiques of the methodology adopted within that study were:

- Questions that have been copied-in can reduce the perception of interviewer attentiveness.
- Interviewees can feel that they are not being listened to.
- Interviewees did not spend sufficient time reinstating the context or reviewing the summary.
- Pressure on the interviewer can result in misalignment and errors, especially with respect to the evaluative nature of verbal newsmarks and continuers

The following broad changes to the technique are suggested in order to address these shortcomings.

- That the interviewee is explained that some standard questions will be copied-in, and that the reason for this is so the interviewer can spend more time understanding and reading their responses.
- Interviewees are told that it is normal to feel unsure as to whether the interviewer has listened to you and that it can take some time for the interviewer to respond.
- Interviewees will be given a fixed time to reinstate the context and summarise. The interviewer will contact them after this time period has elapsed.
- Pressure will be taken off of the interviewer by modifying their client to easily facilitate the copying of prefabricated question stems, note taking, and time-keeping for each phase.

The specific interviewing schedule to be used in this study will be outlined in the methodology section.

### 8.1.2 The Content Criteria

The content analysis technique of CBCA works on the transcripts generated within an interview. Within CMC, these transcripts are recorded by the software, but there also exists an opportunity to examine something that is not possible face-to-face; the deleted text. It is difficult to assess how to treat the utterance segments that the interviewee has chosen to delete and possibly retype before a posting is sent to the interviewer. On the one hand, this information may reveal contradictions and different versions of a deceiver's fabrication. On the other hand, even truth-tellers may use the input box in order to 'think aloud'. Coding this information can be tricky as well. For example, does a deleted contradictory piece of evidence indicate deception, or a spontaneous correction indicative of veracity? The suggestion here is that at present the deleted information should not be considered during the content coding process, but a separate form of analysis will be undertaken in this study to examine this information in more depth.

Many of the CBCA criteria would seem to be applicable to the CMC domain as well as to the face-to-face domain. However, there are four of the CBCA criteria that could be affected by the peculiarities of the CMC interviewing technique; Logical Structure, Spontaneous Corrections, Unstructured Production, and Details Characteristic of the Offence.

The history pane provides the interviewee with a form of enhanced memory. This could make the telling of a complex fabrication easier, because it will be easier for a deceiver to track the flow of their story. This affordance is also available to a truth-teller, but because truthful stories are more likely to be near the ceiling of total logical consistency, it seems as though this criterion might lose its discriminatory power. This effect could be reduced if the interaction is fast paced, whereby the interviewee has little time to review the history between utterances. However, fast-paced interaction puts both the interviewer and the interviewee under pressure, perhaps causing them to orient to a formalised turn-taking protocol. This can aggravate a situation where the interviewee already finds it difficult to interrupt. The presence of spontaneous corrections is also linked to the interviewees ability to 'interrupt'. As was shown in the previous study, this can be difficult within CMC and may lessen the discriminatory power of this criterion as well.

Text-based communication shares some characteristics with other written forms. One of these norms is linearity. Business documents, novels, essays written at school, all aim to follow a linear, well-structured style, and it is possible that this norm will transfer to the CMC interview. It is possible that this effect will lessen the discriminatory power of the Unstructured Production criterion.

The nineteenth CBCA criterion refers to characteristic of the account that resembles a particular crime. CMC based crimes are not the topic of this study, but for the content analysis to be used in this way it will be necessary to develop a profile of certain types of CMC crime, such as ontological deceit within the context of paedophilic behaviour.

### 8.1.3 The Nonstrategic Leakage Criteria

In Chapter 5, some of the nonstrategic cues of interpersonal deceit within CMC were hypothesised. These included the way in which deletions were made during the construction of a fabricated account. When a user is fabricating an account, they may modify their utterances considerably before they are posted. This would be reflected in the content of the deleted utterances and the amount of deletions made.

The content of the deleted utterances can be coded according the number of corrections that were made to the actual content of the utterance. These would include name changes or different descriptions (as opposed to simple spelling corrections for example). The quantity of deletions can be assessed by examining:

1. The ratio of 'backspace' key presses to the total number of key presses, **Ratio**  $b:total$
2. The average length of successive 'backspace' strings
3. The ratio of 'backspace' strings to the total number of key presses, **Ratio**  $bstring:total$

In this context, a 'backspace' string is a sequential, uninterrupted sequence of presses of the backspace key.

This information can all be recorded covertly by chat software, however as with all lie-detection cues they are only effective if they are resistant to countermeasures. CBCA is successful because it is supposedly too difficult for most deceivers to construct accounts that appear veracious. The Guilty Knowledge test is successful because deceivers are pressured to respond quickly. Both these elements suggest the possible success of the aforementioned cues and these will be explored further in the current study.

## 8.2 Hypotheses

This study tests the following hypotheses:

1. Truthful accounts generated through a CMC-based interview will achieve significantly higher CBCA scores than fabricated accounts.
2. The CBCA criteria of Unstructured Production, Logical Consistency, and Spontaneous Corrections will have a relatively lower discriminatory power within this scenario.
3. Deceivers undertaking a CMC-based interview will exhibit greater deletion behaviour than truth-tellers.
4. Deceivers undertaking a CMC-based interview will exhibit a greater number of content corrections in their deleted text than truth-tellers.

## 8.3 Method

This study involves an experimental scenario whereby participants were interviewed about two day trips that they had experienced. The participants were each told to tell one true story and one fabricated story, and these formed the two repeated-measures conditions for the experiment. The transcripts were coded using two raters that were blind to the ground-truth of the transcripts.

### 8.3.1 Participants

Twenty-five students (15 male, 10 female, mean age = 21 years 4 months, SD approx. 2 years 5 months) were recruited for the study. The participants self-selected through an advertisement posted throughout the University. They were paid £6 on conclusion of the procedure that took approximately one-and-a-half hours.

### 8.3.2 Design and Protocol

#### *8.3.2.1 Brief*

Upon arrival, the participant was greeted by the interviewer and shown to the briefing room. Here, the participant was given a verbal explanation of the aims of the study, the protocol, confidentiality and anonymity, and payment. The participant was then provided with a written information sheet and a consent form (these have been provided in Appendix IV).

The participant was told that they were going to be interviewed about two day trips that they had experienced. For one of the day trips, the participant was to tell the truth, as best they could remember, but they were to completely fabricate the other account. The order in which the stories were to be told was left to the participant, and they were asked not to reveal which story was which until the study had been concluded. The participant was asked to make their fabricated account as believable as possible, and that they were not allowed to simply re-tell stories with simple modifications (such as by the substitution of dates and times). The participants were told that the interviewing software was recording the transcript, as well as some other data, but participants were not told that the software was recording data on the nonstrategic leakage cues until the end of the study. Upon answering any further queries, the interviewer then showed the participant to the interview room.

The interview room utilised a PC of the same configuration as was used in the previous study. The interview software had been modified for the study and these changes are discussed in the next section. The participant was introduced to this system, and asked to enter a name (these were later replaced by unique participant codes before the transcripts were coded). Once the participant seemed comfortable with the system, the interviewer went to a PC in the adjacent room.

8.3.2.2 The Interviewing Software

The interviewing software used in the previous study was modified to reflect the requirements of the protocol and to facilitate the task of the interviewer. The software comprised of two client programs, one for the interviewee and another for the interviewer.

The Interviewee client was altered to ask the interviewee which story was true and which was fabricated. Upon selecting the Exit button at the top of the main chat window, the participant was shown a dialogue as presented below. The participant would then select which of their stories was the true one by pressing the appropriate button. The client would then save this information to file before presenting the Save Data dialogue (see Figure 8.1) and sending a signal to the interviewers chat window indicating that the selection had been made.

The Save Data dialogue provided the interviewee with a list of which data was to be saved. This data was recorded automatically by the client and was grouped into three distinct types; transcript, keyboard, and scroll bar movement. This facility was provided so that if the interviewee was unhappy about the recording of covert data then they could chose not to have this particular data saved as opposed to pulling out of the study altogether.

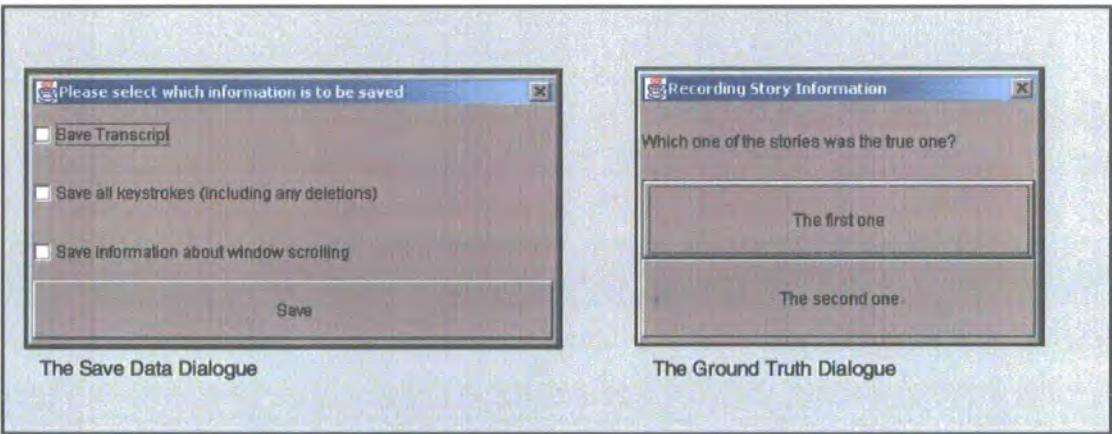


Figure 8.1: The Save Data and Ground Truth Dialogues

The interviewers client was modified to provide an 'Interviewer Assistant' in a separate window. This assistant provided the facility to (1) time each phase of the interview, (2) quickly copy question-stems into the chat client, and (3) provide an area for note taking. The interviewer assistant is shown in Figure 8.2 below.



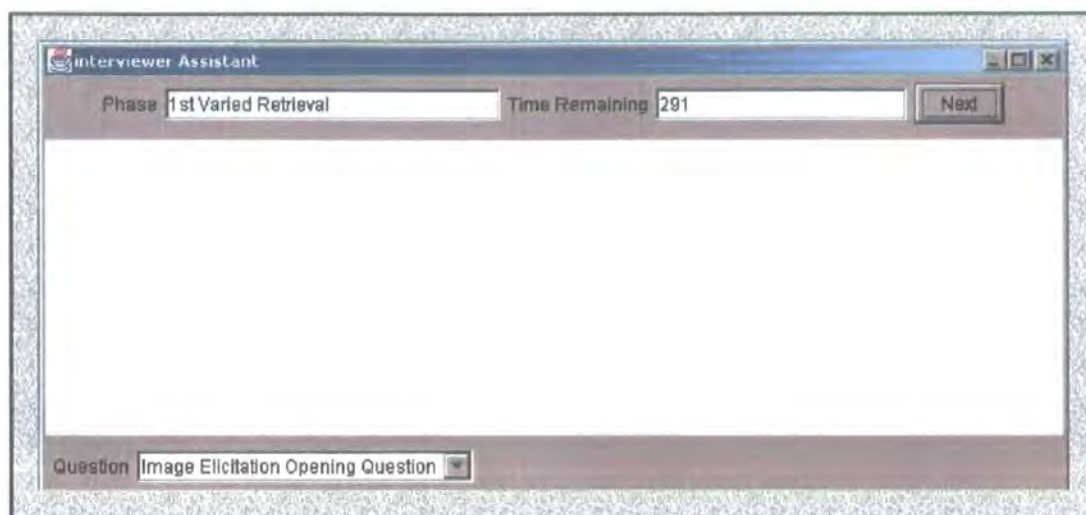


Figure 8.2: The Interviewer Assistant

The scheduler is found at the top of the window. Upon clicking the 'Next' button, the scheduler will jump to the next phase in the interviewing schedule and will provide a count down (in seconds) of the allocated time for this phase. Upon counting down to zero, the scheduler will remain there until the next button is pressed again. Thus, the emphasis is on keeping each phase to its allocated time and to prevent the interviewer from inadvertently 'making-up' time by shortening other phases.

Question stems are chosen selecting a label from the pull down list at the bottom of the page (these stems are provided in the Interview Schedule). Upon selection of the question label, the question stem is copied straight into the chat entry box of the client waiting final editing by the interviewer before it is posted. Notes can be taken in the central portion of the Interviewer Assistant.

### **8.3.2.3 The Interview Schedule**

Some key alterations were made to the interviewing schedule used in the previous chapter. These changes resulted from the suggestions made in section 8.1.1 and the time constraints of the study. The rapport stage was merged with the study brief and was therefore taken offline. In contrast to the procedure adopted in the previous study, interviewees were explained that it was natural to feel a bit anxious about being 'heard'. In addition they were told that the interviewer would be using a cut-and-paste facility, so that the interviewee would not have to wait too long to be given a question. This also meant that the interviewer could post questions more quickly, thus increasing the efficiency of the procedure without exhibiting poor listening skills. The entire interview schedule has been reproduced below.

Phase	Description	Duration
<b>Offline Greeting</b>		
<b>Offline Explanation of Temporal Features of Chat Rooms</b>	Explanation that it is normal to feel a bit anxious about whether or not collocutor has read the message that has been posted.	
<b>Transition to Online Context</b>	Interviewee is brought to Conference Room and instructed in the use of the software. Told that the Interviewer will make first contact through the system.	
<b>Online Greeting</b>	<p>Interviewer greets the interviewee informally with the dialogue.</p> <p>"Hi there, can you read this?"</p> <p>&lt;response&gt;</p> <p>"Are you comfortable?"</p> <p>&lt;response&gt;</p>	
<b>Loop Twice</b>		
<b>Preparation Time</b>	<p>The Interviewee is given five minutes to prepare their first account:</p> <p>"Okay then, I would now like you to spend five minutes thinking about your first day trip. I will contact you through the system when the five minutes are up."</p> <p>The second account will read.</p> <p>"Great, now I'd like you to think about the second day trip. Again, I will give you five minutes to do this and will contact you once those five minutes are up."</p>	5 Minutes
<b>Initial Question</b>	"Okay then, so what trip would you like to talk about [first of all]/[now]"	
<b>Contextualisation Question about preceding events</b>	An Initial question that acts to put the event in context. For example "So why did you decide to go Brighton that weekend?"	5 Minutes
<b>Request for an Extended Narrative</b>	"Okay, so tell me about the whole trip from start to finish. Remember, I know nothing about the experience so I want you to take control of the conversation. If at any point you wish to add something then feel free to cut in. Additionally I'd like you to remember as much as you possibly can about the experience. Okay! So with that in mind, tell me about what happened?"	10 Minutes
<b>Questioning about three selected images.</b>	The extended narrative will make reference to people, places and events that are to be probed further in this section. Three such items will be focused on, and will be probed by first asking the interviewee to concentrate on the	10 Minutes

	remembered image, then by asking an initial open question about the image which will be followed by a sequence of short answer questions.	
<b>Varied Retrieval on fourth image (perspective of another person)</b>	Another image from the narrative should be selected that involved another person. The interviewee is then asked to relate the image from the perspective of that person (i.e. to role play).	5 Minutes
<b>Summary</b>	The interviewee is asked to review the chat programs history window for amendments or additions.  "Now I'd like you to look over the transcript in the history window. Is there anything of interest that you'd like to correct or add about the trip."	
<b>Online closure</b>	The interview is terminated with the following dialogue.  "Well that was great, thanks a lot! The interview is now at an end and I will join you from next door. Before then, however I'd like you to click on the 'Exit' button at the top of the chat window and inform the computer about which story was the true one."	
<b>Offline Rendezvous</b>	The interviewer rejoins the interviewee	

The full ECI procedure was conducted twice, once for each of the two stories. The interviewees were referred to by name and not by pseudonym in order to promote personalisation. The questioning phase was fixed at the elicitation of only three images and there was to be one attempt at varied retrieval. This restriction aimed to increase standardisation across the interviews at the cost of constricting the interviewer. The lesser time spent on varied retrieval reflected its lower efficiency in generating new information in comparison to the time it took to conduct. Whilst it is premature to conclude that the varied retrieval phase is too difficult to carry out online, the time constraints posed by this protocol (thirty minutes for each loop of the ECI) required that more focus be placed on the questioning phase that had demonstrated its fruitfulness in the previous study.

**8.3.2.4 Debrief**

At the end of the interview, the interviewer rejoined the participant in the interview room. The software then presented a checklist that included each major category of data that could be saved. The participant was then asked which data they would permit to be stored and subsequently signed a second consent form to that effect.

Following this, the participant was shown back into the briefing room and was reminded not to reveal which story was true and which one was false. The participant was then asked about

how they felt about the experiment and if they had any more questions. Finally, they were reminded of their right to withdraw and were paid the sum of £6.

### **8.3.3 Transcript Coding**

The CBCA was conducted by two independent raters who were unaware of the truth status of the transcripts. The raters were trained<sup>1</sup> by a British CBCA expert and read all of the major published papers on CBCA. Following tuition on CBCA scoring, the expert and the raters scored several example transcripts from a separate study (where the raters were ignorant of the experimental conditions<sup>2</sup>). The results from these transcripts were then compared and feedback was given to the raters.

The transcripts were stripped of all personally identifying material. Only the first fifteen of the CBCA criteria were used in the analysis. The criteria c10, c17, c18, and c19 were not used, as they were oriented specifically towards sexual abuse crimes. Each criterion was rated on a five-point scale that ranged from 'Absent' to 'Strongly Present'. In addition, each criterion (except c1 and c2) was scored for frequency. A more subjective scale was also used for the rater to express a judgment concerning the overall veracity of the transcript. This bipolar five-point scale ranged from 'Completely False' to 'Completely True'. The raters also reported their confidence of this judgment on a five-point scale that ranged from 'Not at all Confident' to 'Completely Confident'.

## **8.4 Results**

### **8.4.1 Tactics of Deceit**

During the debrief, the participants were asked about the tactics they had used in order to make themselves seem more convincing. A list of these has been provided below.

1. When asked about people they imagined other people and drew from these descriptions.
2. Put in as many facts as possible.
3. Put in details that were mundane or 'by-the-by'.
4. Attempted to keep a consistent style (pace and grammar) between the truthful and fabricated stories.
5. Guessed what close friends would say or do.
6. Used an outstanding event.
7. Mentioned talk and situations as opposed to mere actions.
8. Drew from collated and merged experiences.
9. Drew from books.

---

<sup>1</sup> Raters were trained and provided by the Psychology Department of Portsmouth University. This description of the rater-training programme was provided by Dr. Lucy Akehurst (2003: personal correspondence).

10. Were wary of scrolling behaviour.
11. Tried to maintain the same level of detail between the truthful and fabricated stories.
12. Mentioned affect as opposed to mere events.
13. Gave spontaneous details at the end.
14. Used verbal signs of thinking (e.g. Umm).
15. Tried to be consistent and to refer back to previous details.
16. Tried not to be too definite or confident.
17. Admitted lack of memory.

What is striking about many of these tactics is that they attack many of the CBCA criteria. For example, tactics 2 and 3 confound the 'Quantity of Details' criterion, tactics 16 and 17 confound the 'Admissions of Lack of Memory' criterion, and tactic 7 confounds the 'Reproduction of Speech' criterion. The most common tactic was to draw from real-life experiences and this would result in fabrications that were not strictly 'bare-faced lies'. However, this tactic is likely to be used in naturalistic scenarios and so it is important that CBCA should be resistant to such strategies. However, participants never reported using more than two or three of these tactics at once, and so CBCA's holistic approach may yet be robust to such countermeasures.

## **8.4.2 CBCA Scores**

### ***8.4.2.1 Overall Judgements of Veracity***

The subjective ratings of veracity were assessed for each of the two raters. Rater 1 achieved a total accuracy rate of 66% with 76% of the truthful and 56% of the fabricated stories being correctly rated as such. From a truth-validation perspective this lead to a false-positive rate of 22% and a false-negative rate of 12%. Rater 2 showed a lower total accuracy of 48%, resulting from a significant truth-bias. Here 80% of the truthful stories were correctly identified, but only 16% of the fabricated stories (leading to a large false-positive rate of 42% and a false-negative rate of 10%).

Raters agreed in only 66% of their judgments and there did not seem to be a relationship between agreement and ground-truth (45% of the agreed transcripts were fabricated and 55% were truthful). Only 60% of the agreed transcripts yielded accurate ratings by the raters. The low-level of inter-rater reliability exhibited in the overall judgments was also reflected in the strength of presence ratings of each of the individual criteria. This is discussed further in the next section.

Alongside their overall veracity judgments for each transcript, raters were also asked to assess the confidence of their decisions, along with the degree to which they believed that the

---

<sup>2</sup> The raters were aware that some of the transcripts would be truthful and some would be fabricated.

statement was either completely true or completely false. The means of these ratings for each rater across the truthful and fabricated transcripts are shown in Table 8.1 and Figure 8.3.

Rater	Transcript	Average Total				Average Correct				Average Incorrect			
		Deg.		Conf.		Deg.		Conf.		Deg.		Conf.	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
1	Truthful	3.32	0.69	2.88	0.83	3.58	0.51	2.84	0.83	2.50	0.55	3.00	0.89
	Fabricated	3.16	0.85	2.76	0.78	2.42	0.51	2.75	0.87	3.85	0.38	2.77	0.73
2	Truthful	3.75	0.74	3.00	0.98	3.95	0.60	2.70	0.73	2.75	0.50	4.50	0.58
	Fabricated	3.64	0.49	3.20	0.96	3.00	0.00	4.00	0.00	3.70	0.47	3.13	0.97

Table 8.1: Degree of Veracity and Confidence Ratings

The plot in Figure 8.3 illustrates important features of this data. Firstly, a truth-bias is represented by the clustering of total degree scores in the truthful partition. Secondly, it can be seen that either rater rarely scored a transcript at the extreme poles of the confidence or degree scales. Thirdly, Rater 2 showed considerably more confidence than Rater 1 in their decisions over the fabricated transcripts and this difference contrasts with the similar scores attained for the truthful transcripts.

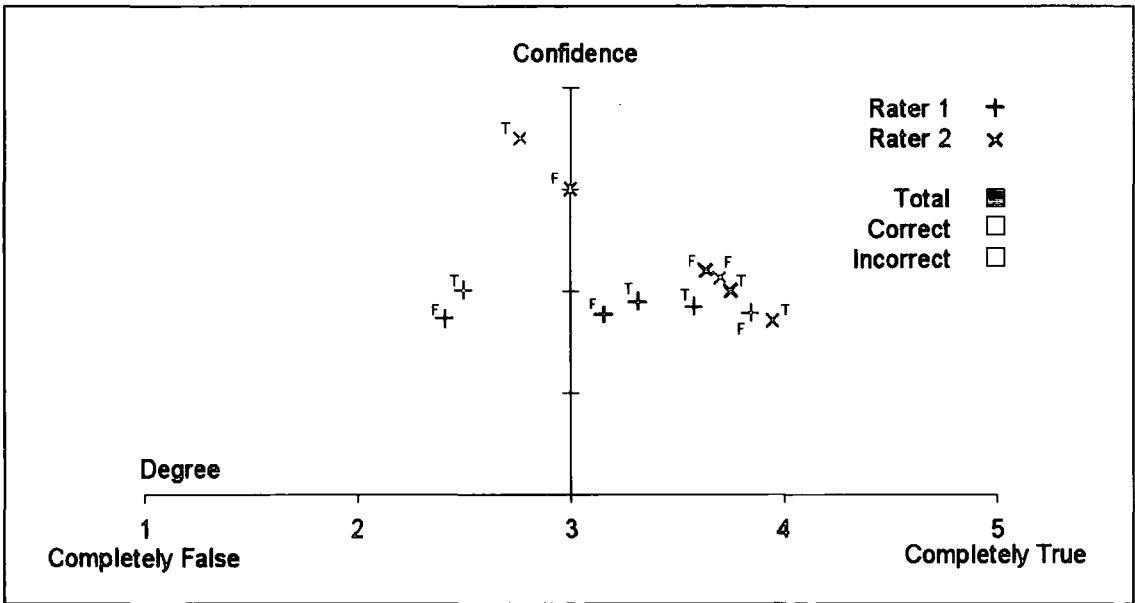


Figure 8.3: Plot of Degree Against Confidence

#### 8.4.2.2 Inter-rater Reliability

Inter-rater reliability was calculated for both the frequency scores and the more subjective scores for criteria presence. The latter of these are presented in Table 8.2 below. The

percentages designated as 'Strict' represent the number of times across the fifty transcripts that the two raters gave the same score for a particular criteria. The results showed a poor level of inter-rater reliability across all of the criteria.

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15
<b>Strict</b>	26	38	38	60	24	84	38	48	88	28	62	60	76	44	78
<b>Recoded</b>	72	48	64	76	56	94	74	60	98	36	64	66	98	52	100
<b>Weak</b>	90	86	92	100	68	100	80	92	100	80	98	100	100	94	100

Table 8.2: Inter-rater Reliability

The table shows that under the strict classification only four of the criteria managed to attain a reliability score of 70% and none achieved greater than 90%. It was hypothesised that these poor figures could be explained by a difficulty in conducting a subjective assessment along a five-point scale. The scores were subsequently recoded along a three-point scale whereby scores of 1 and 2 were amalgamated, as were scores of 4 and 5. The new inter-rater reliability scores have been presented under the 'Recoded' classification in Table 8.2. The recoding procedure yielded some improvement in reliability scores but this improvement still did not yield acceptable results. Furthermore, the criteria that did receive high reliability scores were those that were usually rated as 'Absent' across all of the transcripts (see section 8.4.2.3).

Further analysis of the data revealed that the differences between rater judgments were usually no more than a single point on the scale. Nonetheless, many of these judgments crossed the boundaries of the recoded three-point scale and so this recoding scheme may not actually have been appropriate (also this scheme has a bias towards bipolar scales). Inter-rater reliability was then re-scored using a 'Weak' classification scheme that allowed differences of no more than one point on the scale. These results are presented in the final row of Table 8.2. It can be seen that the results are significantly improved under the weak classification scheme, implying that although raters did not agree often, when they disagreed it was usually by no more than a single point on a rather subjective scale. In the next section, multivariate analyses are reported that utilise the subjective scores of the eleven criteria that satisfied 90% reliability levels on the weak classification scheme. The scores for both raters were combined to form mean values.

A cursory glance of the frequencies scored by both raters showed a very poor level of inter-rater reliability under a strict classification. As these scores were supposed to be indicative of a more objective quantity, it did not seem appropriate to attempt to combine the scores of each rater into a mean value (it is unclear what such a score would represent). To investigate the results in more depth, a multivariate analysis was conducted on the frequency scores of each rater independently under the assumption that any effect would only at best lend very limited support to Hypotheses 1 and 2.

8.4.2.3 Multivariate Analysis of Individual Criteria

Three separate multivariate analyses were conducted to assess the effectiveness of the criteria in distinguishing between truthful and fabricated transcripts. Each analysis concerned one of the three variables groups mentioned in the previous section, (1) the mean subjective criteria-presence scores, (2) the frequency scores for Rater 1, and (3) the frequency scores for Rater 2.

A MANOVA was conducted on the mean subjective criteria presence scores with ground-truth as a factor. Eleven criteria were used in the analysis (c1, c3, c4, c6, c8, c9, c11, c12, c13, c14, c15) and no significant effect was found,  $F_{(11,38)} = 0.72$ ,  $p = 0.712$ . The univariate effects are shown in Table 8.3 for each of the individual criteria.

	c1	c3	c4	c6	c8	c9	c11	c12	c13	c14	c15
Truth	3.92 (0.59)	4.00 (0.80)	3.56 (0.63)	1.50 (0.69)	2.06 (0.49)	1.12 (0.30)	2.90 (0.74)	2.68 (0.78)	1.26 (1.41)	2.74 (0.79)	1.16 (0.24)
Fabrication	4.20 (0.38)	3.90 (0.87)	3.56 (0.56)	1.30 (0.50)	1.94 (0.60)	1.20 (0.54)	2.86 (0.73)	2.68 (0.85)	1.14 (0.27)	2.64 (0.62)	1.06 (0.17)
$F_{(1,48)}$	3.97	0.18	0.00	1.37	0.60	0.42	0.04	0.00	1.48	0.25	2.97
p	0.05	0.67	1.00	0.25	0.44	0.52	0.85	1.00	0.23	0.62	0.09

Figures in the first two rows report mean values with standard deviations in brackets.

Table 8.3: Univariate effects of the Subjective Criteria-Presence Scores

Two other MANOVAs using ground-truth as a factor were conducted on the frequency scores of each of the two raters. Frequency data was not coded for c1 and c2 as these criteria pertain to the transcript as a whole. No significant multivariate effects were found for either Rater 1 ( $F_{(13,36)} = 1.557$ ,  $p = 0.145$ ) or Rater 2 ( $F_{(13,36)} = 0.581$ ,  $p = 0.843$ ). The univariate data for both raters have been included in Tables 8.5 and 8.6 below.

	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15
T	184.00 (41.11)	21.44 (6.08)	3.56 (1.83)	0.76 (1.23)	2.12 (1.59)	2.12 (1.30)	0.12 (0.44)	3.40 (1.76)	5.56 (4.32)	4.16 (3.64)	0.36 (0.70)	8.92 (6.84)	0.32 (0.48)
F	179.72 (47.92)	20.92 (6.22)	3.16 (1.11)	0.44 (0.96)	1.40 (0.96)	1.72 (1.43)	0.16 (0.47)	5.52 (2.66)	5.32 (3.36)	3.72 (3.17)	0.16 (0.47)	7.28 (4.10)	0.16 (0.47)
$F_{(1,48)}$	0.11	0.09	0.88	1.05	3.76	1.07	0.10	11.04	0.05	0.21	1.40	1.06	1.42
p	0.74	0.77	0.35	0.31	0.06	0.31	0.76	0.002	0.83	0.65	0.24	0.31	0.24

Figures in the first two rows report mean values with standard deviations in brackets.

Table 8.4: Univariate Analysis of the Frequency Scores for Rater 1



	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15
T	199.12 (51.71)	21.40 (8.08)	0.60 (0.76)	0.92 (1.35)	0.64 (0.86)	1.48 (1.09)	0.16 (0.37)	2.28 (1.79)	6.36 (3.96)	6.16 (3.91)	0.40 (0.91)	5.40 (4.66)	0.00 (0.00)
F	189.32 (55.50)	20.72 (7.19)	0.48 (1.16)	0.52 (1.01)	0.36 (0.57)	1.24 (1.05)	0.24 (0.66)	2.84 (1.93)	5.92 (3.44)	5.72 (4.16)	0.20 (0.50)	4.52 (3.16)	0.00 (0.00)
$F_{(1,48)}$	0.42	0.10	0.19	1.41	1.84	0.63	0.28	1.13	0.18	0.15	0.92	0.61	*
p	0.52	0.76	0.67	0.24	0.18	0.43	0.60	0.29	0.68	0.70	0.34	0.44	*

Figures in the first two rows report mean values with standard deviations in brackets.

\* No statistic generated

**Table 8.5: Univariate Analyses of the Frequency Scores for Rater 2**

The only univariate effect of  $p < 0.01$  was that of c10 for Rater 1. However, a comparison of the means shows an effect contrary to the directional hypothesis that truthful statements will show a greater frequency of criteria than fabricated ones.

These results tend to confirm the null hypothesis for Hypothesis 1. This conclusion may be qualified by taking in consideration the poor inter-rater reliability, but nonetheless little support has been provided that demonstrates a higher presence of CBCA criteria in truthful statements within this scenario. A fuller examination of why this might be the case is provided in section 8.5.

### 8.4.3 Deletions

Deletion data was examined for twenty-four of the participants (one participant accidentally selected not to have their data saved). A two-tailed matched pairs t-test revealed no significant differences in Ratio  $b_{total}$  ( $t = -0.425$ ,  $df = 23$ ,  $p = 0.675$ ), Ratio  $b_{string:total}$  ( $t = -0.085$ ,  $df = 23$ ,  $p = 0.933$ ), or average backspace string length ( $t = -0.405$ ,  $df = 23$ ,  $p = 0.690$ ) between truthful and fabricated accounts across the sample. Similarly, the one-tailed prediction that deceivers would exhibit greater deletion behaviour failed to reach significant levels (Hypothesis 3). The individual deletion statistics are given in Appendix IV.

An examination of the contents of deleted fragments also revealed no significant qualitative difference between truthful and fabricated accounts, failing to support Hypothesis 4.

## 8.5 Summary and Discussion

This study has rejected four alternative hypotheses regarding the behaviours of truth-tellers and deceivers within CMC. It was found that SVA in its current form was unable to distinguish between truthful and deceptive statements generated through a computer-mediated ECI. It is important to explore why the scenario failed to reject each of the null hypotheses and this is the purpose of this discussion. As with the previous studies reported within this thesis, a more thorough evaluation of the modus operandi of the research will be deferred until Chapter 9.

Hypothesis 1 and 2 required the use of CBCA to rate the transcripts for veracity. The CBCA used three different types of dependant variable in order to score the criteria; subjective total scores, frequency, and criteria presence scores. Two striking features of the data was the very low level of inter-rater reliability coupled with the low discriminatory power of each of these dependants. These two features are naturally related. The low-level of inter-rater reliability may reflect ambiguity in how the raters interpreted features of the transcript. This difficulty may have resulted in a sufficient amount of noise that interfered with the discriminatory effects of veracity. From this it may be surmised that the training system currently used for raters in the face-to-face domain would need considerable adaptation in order to be effective in the CMC domain.

An alternative explanation for the failure to reject the null hypothesis may stem from the structural features of the CMC interview itself. The transcripts were obtained within a fixed time frame and it is possible that this constraint will have stifled the transcript in such a way that certain criteria (primarily c2) lost their discriminatory power.

A third explanation might stem from the tactics used by deceivers. It was mentioned that many of the tactics discovered would undoubtedly lessen the effectiveness of certain CBCA criteria. The tactics used by deceivers is rarely explored within CBCA and SVA studies, and it would seem that this might be an important area for investigation if the technique is to be used with adults.

It might be the case that it is too difficult to adapt SVA to CMC in this manner, although more research would be needed to confirm that. It may be the case that the technique might be more effective if used with children. At the start of Chapter 3, Paul Ekman was quoted as suggesting that it is too difficult to detect lies through the verbal channel because of the ease at which a person can control what they say. Whilst this statement may not be justifiable in spontaneous face-to-face interaction, perhaps in CMC it carries an extra weight. Within CMC, the heightened impression management processes may enhance a deceiver's ability to employ countermeasures, tactics at which adults may have a considerable advantage over children.

It would seem as though interviewees do not alter their usage of the backspace key when they are fabricating as opposed to when they are telling the truth (Hypothesis 3). In this scenario, deletions rarely changed the content of the utterances in either the truthful or the fabricated statements (Hypothesis 4). One reason for this may stem from the temporal pressures of the online interview, such that the interviewees feel that they do not have much time to edit their responses and compose them mentally before writing them down.

It may also be the case that the textual construction of utterances in QS-CMC does not actually occur as frequently as was first hypothesised. In conversation, little time is spent over mentally revising utterances, and even though the QS-CMC environment may put the interviewees under more pressure to respond quickly (because the interaction has been slowed) it does not seem as though interviewees feel the need to 'think aloud' by using the chat-entry box.

The heightened cognitive effort that was expected to result in more deletions (such as spelling corrections) during deceit also did not occur. It may be that this type of lie was not as difficult for the interviewees as might first have been assumed.

Finally, it is important to note that the external validity of this study may be restricted. The scenario of using a day trip as the subject matter is in many ways different to the scenarios that are often used when testing CBCA. However, this study has shown that within this particular scenario the technique has failed to distinguish between truthful and fabricated eyewitness accounts. This finding highlights important limits to the robustness and range of deceptive phenomena that could be assessed by a CMC-based SVA.

## Chapter 9

### Evaluation

---

This thesis set out to examine interpersonal deceit within CMC and to take a first step towards the development of lie-detection techniques that use the medium. To this effect, the literature survey in Chapters 2, 3, and 4 reviewed the work that has been conducted in the areas of interpersonal deceit, lie-detection, and CMC. These research fields were then synergised in Chapter 5, where Interpersonal Deception Theory was applied to the mediated domain and qualitative research revealed some of the strategic behaviours to which this theory refers. The experiments reported in Chapters 6, 7, and 8 tested hypotheses that centered upon the application of existing lie-detection techniques to the CMC context. In Chapter 6, some evidence was found that non-strategic leakage clues to deceit were indeed detectable by examining mouse-click response times. The small effect size however limits the applicability of this finding. In Chapter 7, interviewing for the purposes of truth-validation was examined, and it was found that it was possible to conduct a non-corruptive and detailed interview using QS-CMC. In Chapter 8, the technique of Statement Validity Assessment was conducted using transcripts obtained through CMC. It was found that applying the technique to the domain would need more adaptation that was previously thought, as the CBCA was unable to distinguish the truthful from the fabricated statements above the level of chance. It was also found that the deletion behaviours of deceivers within this scenario matched those of truth-tellers to the extent that no discriminatory power (either quantitatively or qualitatively) could be unearthed.

This chapter now proceeds to evaluate the novel research described within this thesis. Each study is examined separately in terms of the validity, realism, and reliability criteria that are deemed most appropriate for each relevant methodological approach. The role that each study contributed towards advancing the field of CMC-oriented lie-detection and related areas will be discussed, and the chapter is concluded with an outlook of what future research might be undertaken based upon the findings from these results.

The empirical work detailed throughout this thesis, is comprised of a variety of different methodologies and the evaluation frameworks used here reflect this diversity. In Chapters 6, 7 and 8, the experimental method was employed using statistical analysis to confirm or reject hypotheses. The results gleaned from these studies have been assessed for reliability, internal validity, external validity, construct validity, mundane realism, psychological realism, and experimental realism (see Aronsen et al., 1999 for a description of each of these under the experimental paradigm).

In Chapter 5, a web-based variant of the questionnaire method was used for the purposes of attaining qualitative data. This piece of qualitative research will be assessed according to construct validity, content validity, and reliability.

## **9.1 Interpersonal Deceit within CMC: A Survey (Chapter 5)**

The primary critique that can be leveled against the validity of the online survey relates to response veridicality. It was speculated in Chapter 5 that there could have existed the temptation for some respondents to fabricate their account of online deception. Whilst this is a weakness of any method that relies upon self-reports, this study would seem to be particularly vulnerable because of respondent anonymity and the nature of the subject matter. Even so, this type of deceit would yield very little observable feedback or reward in comparison with the amount of effort required to fill-out the survey. Whilst response veridicality cannot be guaranteed, it would seem reasonable to suggest that the most of the respondents were answering truthfully.

As with any method that relies upon self-reports, validity can also be affected by the ability of the respondent to recall their experiences. For example, it may have been difficult for some respondents to remember or even understand why they might have become suspicious of a deceit. It is possible that the need to provide an answer to this question would then result in confabulation. However, in many of the accounts respondents seemed happy to keep the answer to this question quite broad, sometimes limiting their answer to a general feeling or intuition. Where the clues to deceit were provided in more detail, these were exemplified with specific instances of behaviour and so it would seem unlikely that the answers were being invented. The responses indicated that questions were interpreted correctly and that there was little evidence of misunderstanding.

Whilst the responses generally exhibited a high level of content validity, extended elaborations were only given by a handful of respondents. Despite this, the survey did attain a broad range of accounts although there was a sample bias towards the 'victims' rather than the 'perpetrators' of deceit.

Within the context of an interview or an open-ended questionnaire of this kind, reliability refers to whether the respondent would provide the same information if asked within another time or setting. There was no longitudinal element to this study, partially because the anonymity of the scenario would have made this difficult. However, the web-based setting allowed the respondents to fill-out the survey in their own time, which may have alleviated some of the pressure to complete quickly.

## **9.2 RT Measures to Assess Guilty-Knowledge (Chapter 6)**

The Guilty-Knowledge test attained a strong degree of internal validity. There did not appear to be any other systematic variance that could have resulted in the cause-effect relationship that

was observed between Guilt and RTs. The same experimenter was used for every participant, using the same scenario, and the same replayed transcript. There was a slight variation in how the mock-crime was realised by each participant but no *systematic* difference was found. Any individual differences in reflex-time were compensated for by the repeated-measures manipulation of the study. The order in which participants took the two tests was counterbalanced between participants. A two-tailed t-test revealed no significant main effects for order or gender across the guilty probe, guilty irrelevant and pooled innocent irrelevant items for both RTs and Accuracy (see Table 9.1).

	Gender		Order	
	RT	Accuracy	RT	Accuracy
Guilty Probe	t=2.437, p=0.21	t=-1.135, p=0.266	t=0.675, p=0.505	t=-0.928, p=0.361
Guilty Irrelevant	t=1.931, p=0.064	t=0.414, p=0.682	t=0.084, p=0.933	t=-0.678, p=0.503
Innocent Irrelevant	t=1.039, p=0.308	t=-0.667, p=0.510	t=-0.714, p=0.481	t=1.577, p=0.126

$\alpha = 0.05$ ,  $df = 28$ .

Table 9.1: Ordering and Gender Effects

The primary threat to internal validity came from an error in the experimental protocol. Ideally, the stimuli set used for the guilty condition and the set used for the innocent condition would have been swapped over for half of the participants. This would have allowed a test to see if any between-subject effects resulted from the use of a particular stimuli set. An error in the protocol meant that the same set was always used for the guilty and innocent tests (and subsequently the mock crime). Nonetheless, care was taken in the production of the two sets of stimuli, and it is difficult to see any relevant difference between the two sets that might have resulted in a systematic difference.

The primary threat to the external validity of this study might result from the use of undergraduate students. Much faster response-times were attained in this study than in the studies by Farwell and Donchin (1991) and Seymour et al. (2000). It was suggested in Chapter 6, that this difference could have been a result of the apparatus, but it is also possible that the above-average skills of the participants will have resulted in quicker RTs. This overall effect may have reduced the difference in RTs between guilty probe and irrelevant items. It is possible that the skill of the participant is a factor that can either enhance or diminish the main RT effect of Guilt. Another lesser concern is that the scenario did not involve a punishment or reward element for successful deceit. However, as the guilty-knowledge paradigm does not theoretically depend upon such a device, this limitation was deemed acceptable.

Care was taken with the construction of the stimulus set (the independent variable) so that no systematic differences would result across the different stimuli categories. This enhanced construct validity by ensuring that the irrelevant items *were* actually irrelevant. For each probe item of a particular 'type' there were four other irrelevant items of a similar type. This reduced

the salience of the probe items against the irrelevant items. Selection of irrelevant stimuli is a recognised challenge of the Guilty-Knowledge test (Ben-Shakhar and Elaad, 2002), and it can be difficult to select irrelevant items that have no personal relevance to the participant but are still similar to the crime items. It was believed that the between-measures manipulation of the study compensated for any *individual* differences in the perceived salience of irrelevant items.

The protocol attained a high degree of experimental realism. The mock crime had to be performed under pressure and required participant interaction rather than mere observation. The participants indicated in post hoc interviews that they found the mock crime to be involving and that they believed in the experimental deception. The test itself required a high level of concentration and this was reflected in the high accuracy levels reported in the results.

It is likely that the task being undertaken during the mock crime would have had low mundane realism for the participants. This would have been mediated to some degree if the participants were familiar with chat room systems. Nonetheless, it cannot be assumed what degree of experience the participants may have had with chat room deception, and the test itself was a highly artificial scenario. However, as a controlled lie-detection technique in many ways depends upon the scenario being highly controlled and artificial, the low mundane realism attained by the study was not a concern over-and-above the sample biases already addressed under external validity.

The key psychological realism issue concerned the degree to which the main effect for Guilt was in fact related to the elicitation of the orienting reflex. As no event-related potentials were recorded within this study, no conclusions can be made as to the effectiveness of eliciting this response. However, a main latency effect was found and so it is evident that some sort of recognition reflex must have caused the inhibition of motor control. The most likely candidate for this would seem to be the OR.

The false-negative rate attained by this study gives some indication of the reliability of the latency effect phenomena. Despite showing significant discriminatory power for 76.7% of guilty suspects, it must also be remembered that six of the participants showed no significant latency effects, and one participant actually showed quicker guilty probe RTs. A closer scrutiny of the data for this participant failed to reveal any characteristics that could account for this. Considering the level of control employed throughout the study, the reasons for the low discriminatory power for these participants is most likely to be attributed to countermeasures. Despite the finding by Seymour et al. (2000) that countermeasures did not seem to work, it must be remembered that these investigators also found longer latency effects. It could be possible that shorter overall response times result in lower discriminatory power and the higher potency of countermeasures within this version of the technique. Whether or not this is actually the case remains a question for further investigation.

### 9.3 A CMC-Based SVA (Chapters 7 and 8)

It is important to draw the distinction between an evaluation of the effectiveness of CMC-based SVA and the evaluation of the two studies that have sought to determine that effectiveness. The latter of these two questions will be addressed within this section<sup>1</sup>.

With respect to construct validity, it is possible to identify two main types of construct, the independent (ground truth of the transcripts) and the dependant (the CBCA scores and deletion metrics). As with the study reported in Chapter 5, there is no real way of ensuring that the participants were in fact lying when they were told to do so. It was hoped that the absence of any tangible reward for deceit would deter this form of malicious action. More pertinently, the use of countermeasures by the participant may have blurred a division between truthful and deceptive transcripts in a way that confounds CBCA. However, this is not a fault of the experimental scenario insomuch as it is a very pertinent issue relating to the external validity of the technique.

The main issue of external validity relates to the extent to which the results detailed in Chapters 7 and 8 can be generalised to interviews conducted by other interviewers, possibly using different methodologies or technologies. The results are likely to be restricted to three key features of the domain. The participants were university-educated adults, the interviews were carried out using QS-CMC, and the enhanced cognitive interview formed the basis of the methodology. It is plausible that an alteration to any one of these features may enhance the discriminatory power of the dependants.

Internal validity was ensured by allowing the participants to choose the order in which they told either the truthful or the fabricated story. This reduced the possibility of interviewer effects that would favour one type of statement over another. The within-subject manipulation ensured that the sample from which the fabricated statements were generated was identical to that of the truthful statements.

With regard to psychological realism, the most salient concerns surrounded countermeasures (mentioned previously) and the relationship between the cognitive load and the deceptive scenario. The hypotheses that related to deletion behaviours were founded upon the assumption that the performance of deceit would result in an increase in cognitive load. Cognitive load was not measured independently in this study (as is rarely the case in lie-detection experiments) and so it is difficult to assess whether the participants were finding the fabricating task more difficult than telling the truth. It could even be suggested that fabrication would require less effort within this scenario than remembering a day trip from a time in the past. As with the guilty-knowledge test attempted in Chapter 6, there was no fear of punishment to the deceiver within this scenario. Whilst the inchoate theoretical grounding of CBCA makes it difficult to evaluate what effect this detection apprehension would have on CBCA scores, it

---

<sup>1</sup> The former question was covered in the discussion sections of Chapters 7 and 8.



nonetheless represents a valid distinction between experimental scenarios such as this one, and the empirical roots of SVA.

Two questions can be posed that relate to mundane realism; to what extent does the CMC-based investigative scenario match that of real life interrogations, and to what extent does the adapted ECI engage the participant in a manner that resembles typical CMC discourse. The first question bears little relevance within this evaluation because the technique being proposed by the study is intended to be novel. For the second question, it was shown by the pilot study in Chapter 7 that the style of discourse departed substantially from that of natural CMC collocation. More specifically, the turn-taking system exhibits a more formalised orientation and it is rare that a CMC user would find themselves being asked to provide a free narrative. Therefore, there may exist the possibility that it is inappropriate to generalise the findings from other types of CMC discourse to the ECI scenario.

The interviewees in both the pilot and main SVA study seemed to be engaged in the task. The post-hoc interviews with the participants used in the SVA study, alongside the 'perceived duration' estimates obtained in the pilot would suggest that the participants exhibited a comparable degree of immersion in the activity as is typical in other forms of QS-CMC discourse. In this manner, it would seem as though the studies achieved a high level of experimental realism.

The poor inter-rater reliability results were already discussed in Chapter 8, and this highlights the need for a training programme to be developed that attempts to more carefully fashion the CBCA technique to the CMC domain. The reliability of the deletion data was high and the participants rarely deviated from a certain level of deletion behaviour, regardless of the nature of the story that they were telling. It remains a problem for future research to ascertain to what extent this artifact of behaviour is indicative of the interactive mode (i.e. an enhanced cognitive mode) rather than the level of veracity involved.

## 9.4 Revisiting the Research Questions

In Chapter 1, two broad research aims were presented. The research presented within this thesis has met the objectives that realise these aims. These are now discussed in turn.

### **How do deceivers perform their interpersonal deception strategies whilst communicating through computer-based quasi-synchronous conversation?**

In Chapter 5, the role of context interactivity in IDT was revised and a theoretical application of Interpersonal Deception Theory to the realm of CMC was presented. Qualitative research was also undertaken that described a set of different categories of online deceit and revealed the importance of stereotyping in driving suspicion.

The modest latency effects revealed by the Guilty-Knowledge test and the findings concerning the diagnostic power of deletion behaviours, temper the hypothesis that there are any significant non-strategic deceptive displays that promulgate within the CMC medium. It seems reasonable

to propose that there is a dearth of non-strategic information revealed through CMC and that future research should be conducted under this assumption. This proposal is strengthened by the findings of SIP and the observations made in Chapter 7 regarding interpersonal reciprocity. If collocutors are working to orient to one another's behaviour and to attain relational information, then this would militate the superordinate effect of interpersonal reciprocity (posited by IDT) and subsequently lessen the significance of non-strategic displays.

However, it might also be proposed that collocutors will use verbal cues (stereotypical or otherwise) instead of nonverbal cues to detect non-strategic behaviour. Again, this might occur because collocutors are driven by relational goals and are trying to seek out social information. The findings of SIP would suggest that over time collocutors would become more adept at using these cues, thus increasing the availability of this type of non-strategic information. This is interesting because IDT suggests that the role of non-strategic cues will lessen over time, but if collocutors are utilising verbal cues instead then the opposite might be true.

The victims of deceit surveyed in Chapter 5 revealed that their suspicions were largely based upon the observation of deviations from stereotypical behaviour by the sender. Whilst such strategies have been reported within the face-to-face domain, it is suggested here that within CMC, receivers have a greater dependency on these social schemas due to the absence of any other *non-strategic* information. When the findings of the SIDE model are considered, then the interrelationship between the receiver and a sender within a deceptive episode can become quite complex. Essentially there are at least four main dimensions to this relationship; group membership (ingroup or outgroup), identity salience, context interactivity (see the discussion in Chapter 5), and role (sender or receiver). For example, where individual identity is made salient, and the sender is from an outgroup, then an ingroup receiver becomes predisposed to a greater level of suspicion because:

1. SIDE predicts that the receiver will orient to behaviour that is more personally-relevant.
2. This will increase the receiver's perception of the sender as belonging to an outgroup, resulting in a stereotypical impression.
3. As the sender is also orienting to personally relevant behaviour then they are likely to deviate from stereotypes held by the receiver.
4. This will then result in greater suspicion by the receiver as the sender's behaviour deviates from expected norms.

Contrastingly, suspicion is likely to be considerably lower when both the sender and receiver are from the ingroup and group identity is made salient. The orientation to group stereotypes within CMC will then act to mediate the utility of a receiver's primary cue to suspicion (that of atypical behaviour). This could mean that children could be more at risk from paedophiles engaging in ontological deceit in chat rooms where group salience is high (such as a fan-oriented site), where paedophiles engaging in deceit (based upon *stereotypical* notions of

childlike behaviour) are perceived as being less suspicious by the other users. Fortunately, a moderator (another receiver) would not be as prone to these effects as they would exist outside of the ingroup and are likely to have an enhanced level of context interactivity. The complex relationship between these four dimensions needs further analysis to establish to whether these factors operate subordinate or superordinate to the interpersonal reciprocity dimensions mentioned earlier.

The second research question posed in this thesis referred to the development of a potential CMC lie-detection technique.

**Is it possible to create a lie-detection technique that can distinguish deceptive behaviour reliably from the behaviour of truth-tellers in a controlled CMC-based setting?**

A guilty-knowledge test was successfully conducted using a standard PC. However, the effect size attained in this experiment suggested that detecting deceit at this level might be too vulnerable to countermeasures. The examination of another non-strategic behaviour, deletion behaviours also did not provide a very promising route to lie-detection.

A modified SVA procedure was unable to discriminate between truthful and fabricated account using verbal clues. It was found that participants were able to use some countermeasures that would affect the CBCA scores (although only affecting three of the criteria at most). A SVA procedure, driven through CMC might require a more substantial adaptation than was initially thought. It is suggested that the reliability and validity of SVA might be improved by:

1. Promoting a less structured form of expression through CMC.
2. Refining the content analysis techniques to the peculiarities of CMC.
3. Augmenting the technique by using nonverbal information that is *not* passed through the computer interface.

The CMC-based investigative interviews detailed in Chapters 7 and 8 seemed to be quite successful. Now, more research needs to be undertaken to increase the external validity of the technique, using forensic-style scenarios and a wider variety of interviewers and participants.

Ways need to be found to promote a little more freedom from form. It has been suggested that the use of props (in child abuse cases) can be corruptive and maybe this is the case for computer-assisted props. However, with adults this might not always be the case, and to a limited extent technological props may provide a free-form means of expression that can complement the form of the CMC structured interview. It is suggested here that the CMC-based investigative interview enhances the reliability and validity of statements obtained for use with SVA because it removes many of the interviewer effects that can impede their elicitation. It also helps to add more control to the procedure with subsequently helps to standardise the CBCA technique by more carefully constraining the scope of phenomena used under the analysis.

1. The dislocation and occlusion of nonverbal cues within CMC requires that interviewees have to adapt to the medium, orienting to a more formal style of discourse. This can remove the elements of spontaneity that are critical to the successful investigative interview.
2. This may be helped by augmenting the technique with different means of communicating and expressing ideas online. Such as through the use of virtual whiteboards, or three-dimensional models of the crime scene.

The key point behind any of the devices that are developed is that they enhance expressiveness without imposing a structure on the narrative that could lead to confabulation.

The studies in Chapters 6 and 8 examined the potential for using promulgated nonstrategic cues for lie-detection. The results in these scenarios suggest that such techniques might be vulnerable to countermeasures. However, within a controlled setting, the investigator is not bound to only measuring those behaviours that can be recorded through the CMC system. It has been suggested that the use of nonverbal cues can compliment a procedure such as SVA. However, according to IDT the process of interpersonal reciprocity would seem to eschew many of the nonverbal leakage cues to deceit. It is suggested that:

3. Nonverbal clues to behaviours that correlate with deceit are likely to have a higher discriminatory power when using CMC than using face-to-face because:
  - a. The occlusion of the collocutor lessens the superordinate effects of nonverbal reciprocity posited by IDT.
  - b. The eyewitness' attention will be focused on their task and away from their nonverbal impression management.

Because CMC communication occludes nonverbal behaviour, the superordinate effect of nonverbal reciprocity loses its power. As the eyewitness becomes less concerned about their nonverbal impression management, heightened by their absorption into the task at hand. This might increase the degree to which they exhibit nonverbal leakage cues through the face and other parts of the body. This information can be analysed in coordination with the interviewers questions in order to discover any behaviours that correlate with deceit.

The answers to the research questions posed within this thesis, suggest new directions for future research. These will now be discussed in the next section.

## **9.5 Future Research**

Two broad investigative questions guided the research that was presented in this thesis. The first of these examined the realm of online deceit and the second aimed to see whether CMC could be used as a lie-detection tool. In this section, the knowledge attained through this research is used to suggest directions in which these questions can continue to be investigated.

Qualitative research needs to examine two of the most prominent and damaging types of online deceit that were revealed in Chapter 5 (Ontological and Affective). A variation of the Undeutsch

hypothesis may be posed for these two different forms of deceit, that qualitative differences exist between truthful and fabricated performances of self or state of affection. Such differences could be investigated in two ways possibly with the goal of creating a set of criteria similar to that of CBCA or RM. Research might be conducted within a quasi-experimental field setting using discursive techniques or by using in-depth interviews.

The former of these two approaches might be conducted using a social pragmatist approach to discourse analysis (Turnbull, 2003). This recently constructed approach adopts the best traditions of discursive psychology, conversation analysis, and quasi-experimental research to discover conversational performances that are recurrent within certain types of behavioural display (such as deception). An experimental scenario may be construed where dyads are comprised of a deceiver and a receiver and the transcripts analysed to identify any techniques that are used by ontological deceivers through the CMC medium.

In-depth interviews could be conducted face-to-face, but a respondent may feel inhibited due to the identifiability of the setting (especially if they are the 'deceiver'). Online interviews would seem to offer such a solution, whereby the respondent could be invited to use an anonymous email address or bulletin board system to interact with the investigator. However, the process of anonymising the interview can be paradoxical. Anonymous interaction may act to distance the interviewer and interviewee at the early stages of interaction, and this would weaken the relationship and act to lessen any interpersonal benefit that the interviewee would glean from the exchange. Conversely, anonymous communication may eventually lead to more frank and open discourse (under certain conditions). Whereas the discursive approach may be able to describe how ontological and affective episodes are performed, interviews may link these behaviours to the underlying motives and tactics that drive them.

The research detailed within this thesis has demonstrated that a SVA procedure, driven through CMC might require a more substantial adaptation than was initially thought. It is suggested that the reliability and validity of online SVA might be improved by adapting CBCA in light of what conclusions would be found from the aforementioned research. In addition, it is probably better to focus the efforts of a CMC based SVA back towards the testimonies of children as adults are more adept at using countermeasures that confound CBCA.

## 9.6 Conclusion

This thesis has taken the first step towards the development of naturalistic and controlled CMC lie-detection techniques. The research detailed within the last four chapters has initiated an investigation that could eventually lead to the development of detection tools for a variety of deceptive phenomena. Such phenomena could relate to the activities of paedophiles, corporate deception, and the betrayal of intimacy. From the perspective of controlled lie-detection, it has been shown that the CMC-based investigative interview could have the potential to be a very effective means of eliciting testimony from children or taciturn witnesses. This could increase the validity of testimony that might have otherwise been discounted due to potential interviewer

effects. It has also been shown that the development of verbal detection methods has yet to reach maturity, but that the stereotypic behaviours of deception through CMC do not preclude the development of such techniques.

It is important that research conducted along these lines consider two central questions pertaining to how lie-detection might be used in practice. Firstly, it is important to consider the likelihood that CMC-based lie-detection techniques will only ever attain high levels of accuracy gradually, whereby techniques are continuously refined and the deceptive context is understood in greater depth. This implies that for a time, society will have access to lie-detection systems that have accuracy levels that are only moderately successful (say between 70 and 80 percent). The question needs to be asked as to the circumstances under which these techniques should be employed (if at all) and to what extent the results should be contextualised. This question is related to another pertinent issue that even if we did have access to a lie detector that was highly accurate when should this actually be used. As Sissela Bok (1999) has made clear, not all lies are negative, and in fact, many are a very important part of social life.

To illustrate these issues it is worth exploring the case of online paedophilic deception. Let us assume a technique was available that could detect older men masquerading as young girls or boys within a public chat forum. Here it is important to separate two key components of investigation, the lie-detection tools that indicate deception, and how that deception is interpreted and followed up. It is imperative that the constraints of the former guide the latter. If the lie-detection tool has an accuracy rate of only 75%, then the investigators must be aware that ontological lying in itself is not a criminal activity, and that an over reliance on the tool may divert their attention away from the other 25% of cases. Similarly, even if a tool were 100% accurate, the investigators would need to determine how best to use that tool. It needs to be established to what extent 'non-active' paedophiles that utilise CMC deceit, have become 'active' as a result of the relative safety of the medium, acting as a stepping stone to criminal activity. If this is the case, then the existence of a highly accurate lie-detection technique should be heavily publicised, ensuring that the medium is made inhospitable to paedophilic uses. If, however, these paedophiles are simply using the online medium as a bridge to a real world activity they would have engaged in anyway, then the police must be careful not to render their best means of paedophile detection as useless because perpetrators have simply moved their activities elsewhere.

In online intimate relationships, the benefits of lie-detection are also ambiguous. William Blake wrote in his poem, *Auguries of Innocence* that, "a truth that's told with bad intent beats all the lies you can invent". Never is this more the case than in intimate relationships. Relationships in our society involve the gradual build up of trust and disclosure. The enforced levels of veracity that online lie-detection could bring into the early stages of a relationship may impose a culture to which people are unaccustomed.

As of yet, lie-detection techniques have yet to achieve levels of accuracy and reliability that warrant their use in all but the most advisory of settings. This might actually be a fortunate

occurrence, as in many ways the development of these tools should be a subordinate goal to developing our understanding of the intricacies of truth, lies, and deceit in everyday life. People rarely 'just lie' but mix lies of commission and omission with believed truths and self-delusions. Until such processes are more fully understood, both psychologically and morally, the tool of accurate lie detection, even if it were available, should probably not be used.

# Appendix I

## Addendum for Chapter 5

The questionnaire described in Chapter 5 was constructed as a website. In this Appendix, the questions and instructions posted on the website have been reproduced. The website was comprised of five main pages that could be navigated through a menu on the left hand side. These pages consisted of an introduction (Figure I-iii), an instruction page (Figure I-iv), the questionnaire itself (Figure I-v), an exit page (Figure I-vii), and a page introducing the investigator (Figure I-vi). The instructions and questions that were contained in the questionnaire have been reproduced in Figure I-i and Figure I-ii. Section 4 of the questionnaire was not used in the analysis and so has not been reproduced here.

### The Questionnaire

This Questionnaire has been split up into four sections. You may leave out a section if you wish, but please try to answer all of the questions in each section that you do choose to complete.

This questionnaire has been designed to investigate the ways in which people intentionally withhold the truth from one another when speaking in chat rooms. The questionnaire is anonymous and you are asked to not include any information that may be personally identifying (to you or others) in any way. If you do include any such information by accident, then this will be isolated and stripped out after you have submitted the questionnaire.

Most importantly thank you for your help!

Section 1: Please complete this section if you wish to include some basic demographic details

Q1. Could you please tell me your age?

12 years or under  
13 – 17 years  
18+ years  
Unspecified

Q2. What gender are you?

Male  
Female  
Unspecified

Q3. Roughly how much time do you spend using chat rooms per week?

Up to 1 hour  
1 – 5 hours  
5+ hours

(Optional) Please enter your nationality.

Figure I-i: Questions and Instructions Used in the Analysis (Section 1).

175



**Section 2:** Please complete this section if you can remember a time when you thought that a person you were talking to in a chat room was intentionally trying to deceive you. **Try to include as much information as you can for each question.** Please read through all the different parts of the question before you answer. Thanks!

Q4. What was the other person trying to deceive you about?

Q5. How did they try to accomplish this?

Q6. Was there anything they said that made you suspicious, and if so what was it?

Q7. Was there anything else that made you suspicious?

Q8. What sort of impact would you say this deceit had on you?

Q9. Please describe the kind of chat room that you were in at the time.

Q10. Is this a chat room that you visit often, and if so how often?

Q11. Are there any other details about this experience that you would like to add? If so please use the space below.

Thanks for completing this section! If you want to you can now have a look at sections 3 and 4. If not, you can submit the questionnaire by clicking on the "Submit" button at the bottom of the page.

---

**Section 3:** Please complete this section if you can recollect a time when you have intentionally tried to deceive someone in a chat room. **Try to include as much information as you can for each question.** Please read through all the different parts of the question before you answer. Thanks!

Q12. What were you trying to be deceptive about?

Q13. Did you use any techniques to try and make yourself sound more convincing? If so, what were they?

Q14. Was anyone suspicious of you?

Q14a. If somebody was suspicious, how did you know that they were suspicious?

Q14b. If somebody was suspicious, how did you react?

Q15. Did anyone "catch you out"?

Q15a. If they did, what do you think gave you away?

Q16. Are there any other comments about this experience that you would like to add? If so please use the space below.

Thanks for completing section 3. If you want to, you can now complete section 4, but if not then you can submit the questionnaire by clicking on the "Submit" button at the bottom of the page.

Figure I- ii: Questions and Instructions Used in the Analysis (Sections 2 and 3).



Figure I-iii: The Introduction Page

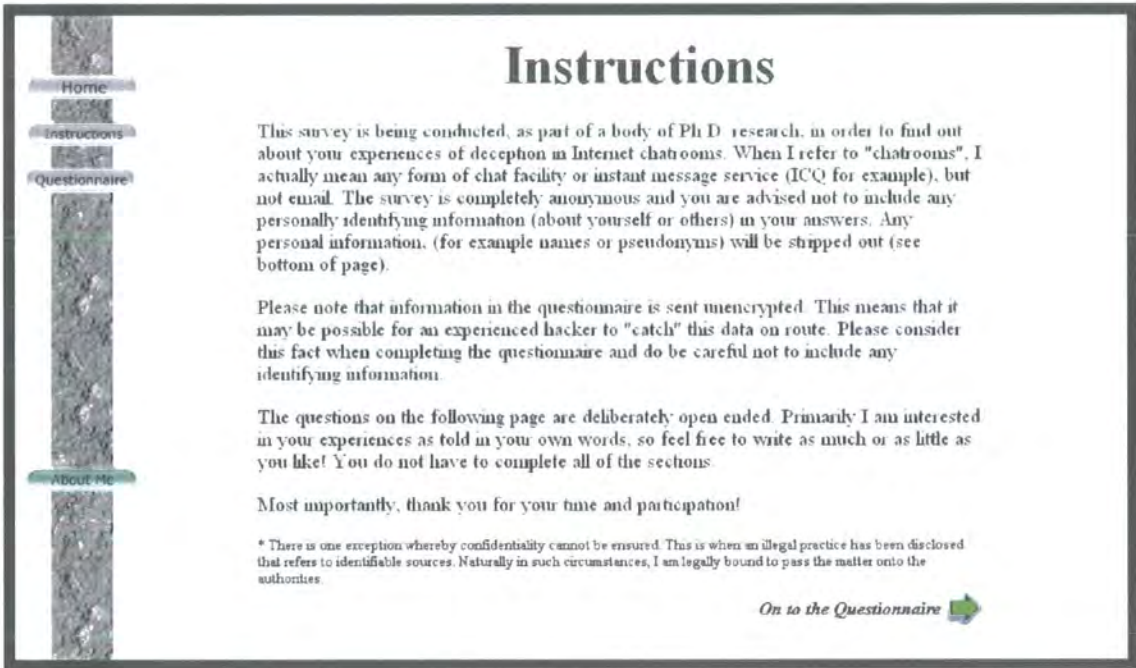


Figure I- iv: The Instruction Page

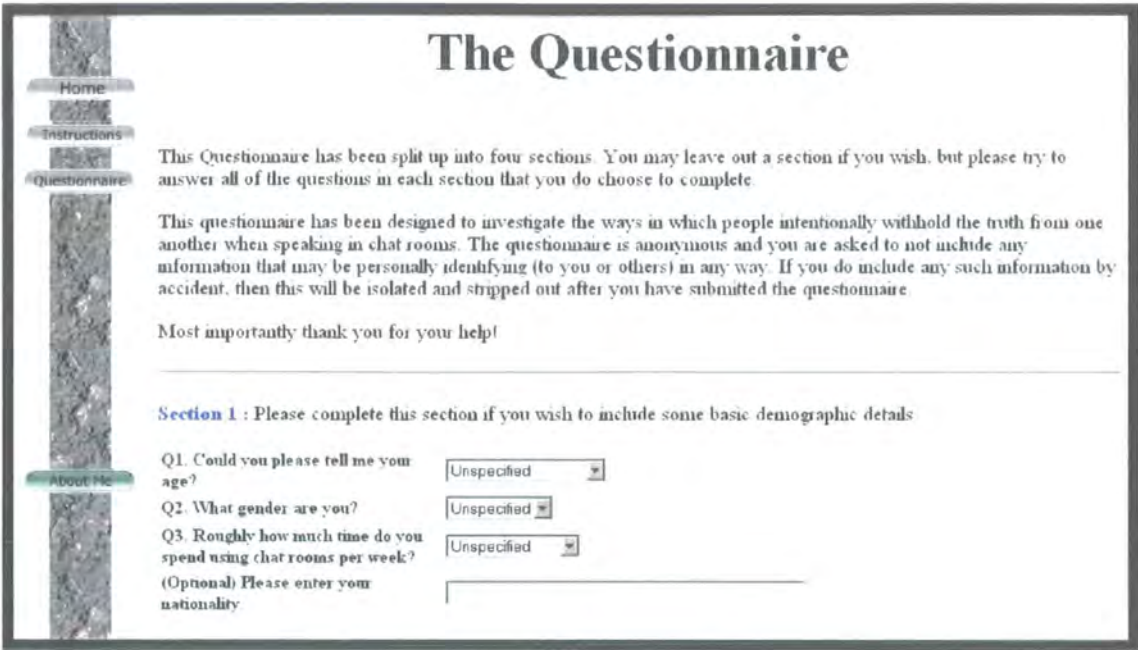


Figure I-v: The Questionnaire Page



Figure I-vi: The About Page

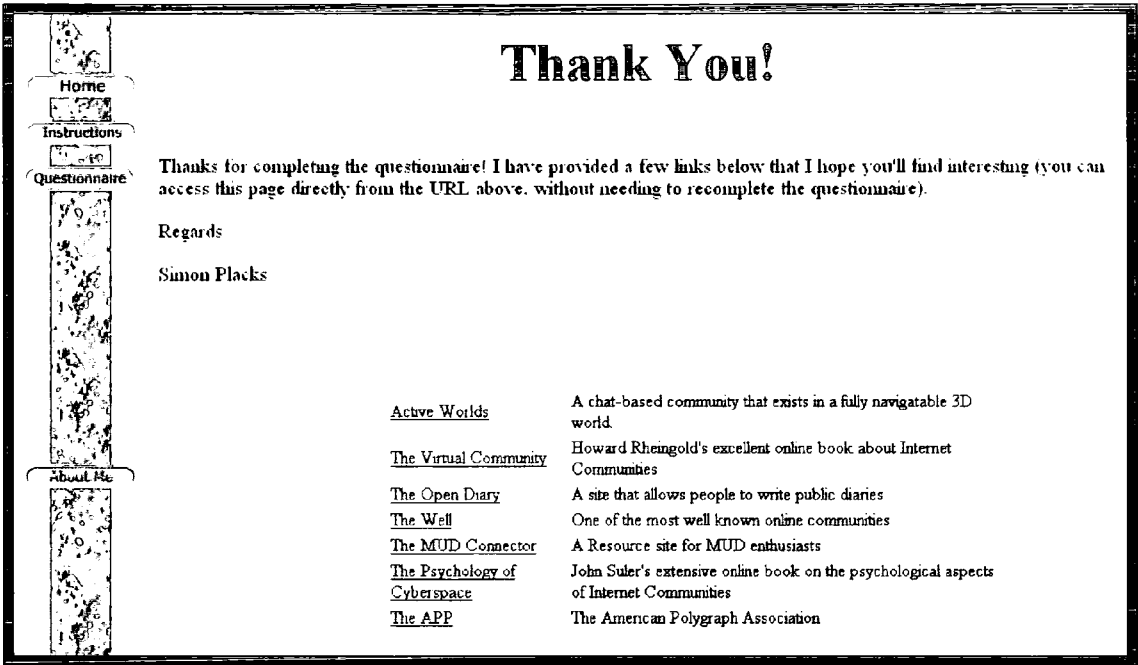


Figure I-vii: The Exit Page

Appendix II

Addendum for Chapter 6

The set of stimulus word pairs used for the Guilty-knowledge Test have been provided in Table II-i. The information sheets given to the participants have been provided in Figures I-i, I-ii, I-iii, and I-iv. Screenshots of the software used in this study have been provided in Figures I-v (the distracter task), I-vi (the testing software), I-vii (the training software), and I-viii (the chat software). A sample of the replayed transcript has been provided in Figure I-ix and the instructions given to the five volunteers who participated in the block-puzzle exercise has been given in Figure I-x.

Guilty			Innocent		
Probe	Target	Irrelevant	Probe	Target	Irrelevant
Greg_Cat	Paul_Hound	John_Mouse	Dale_Hare	Wayne_Bear	Tom_Lamb
		Doug_Wolf			James_Cow
		Mark_Deer			Rich_Horse
		Adam_Dog			Bob_Sheep
Castle_14	Rook_12	Pawn_17	Ace_11	Jack_15	Queen_19
		Bishop_13			Joker_12
		Queen_22			King_20
		King_10			Deuce_18
Block Puzzle	Card Puzzle	Picture Puzzle	Crossword Puzzle	Logic Puzzle	Jigsaw Puzzle
		Domino Puzzle			Word Puzzle
		Number Puzzle			Coin Puzzle
		Marble Puzzle			Match Puzzle
Jane_Star	Kath_Sun	Kate_Comet	Beck_Mars	Rose_Earth	Jean_Saturn
		June_Meteor			Pat_Venus
		Anne_Planet			Ruth_Mercury
		Sue_Moon			Beth_Neptune
Green Cross	Blue Square	Green Triangle	Red Cube	Green Sphere	Blue Tube
		Pink Line			Pink Pyramid
		Yellow Spiral			Yellow Coil
		Purple Oval			Purple Cone
Broken Mouse	Broken Keys	Broken Screen	Broken Button	Broken Switch	Broken Knob
		Broken Disk			Broken Handle
		Broken Memory			Broken Plug
		Broken Speaker			Broken Port

Table II-i: Stimulus Word-Pairs used for the Guilty-Knowledge Test.

## Participant Information Sheet

(You will be given 5 minutes to peruse this sheet)

### 1. Learning Task

**You are now asked to learn the details of the chat room disturbance that you are to commit. The training software will allow you to do this. The main details that need to be learnt are:**

1. The username that you will be known by
2. The password that you will use to log yourself into the system
3. The precise means by which you will disrupt the conversation

**The participants have each been given one piece of a 3-dimensional BLOCK PUZZLE. Their aim is to solve the puzzle verbally, using only the chat room program to communicate.**

You will be using the username **GREG\_CAT**.

Your password will be **CASTLE\_14**.

You will tell the others that your piece of the puzzle forms a **GREEN CROSS** made up of 5 cubes (the solution to the puzzle has no such piece!).

You will then act to disrupt the conversation further by disturbing one of the other members, alias **JANE\_STAR**. You will tell Jane\_Star at least 6 times, that you do not understand what she is saying. You might go about this by using some of the following sentences:

"JANE\_STAR, I don't understand how you came to that "

"JANE\_STAR, I still don't get what your saying"

"JANE\_STAR, what you're saying is not at all clear to me"

**You should ignore her responses and space out your questions in 30-second intervals**

**Remember, your aim is to be as convincing as possible. You should try not to annoy the other members, or make them suspicious in such a way that they end up ignoring you.**

Finally, after 5 minutes have elapsed, you are to type the word **BROKEN MOUSE** repeatedly, about 10 or more times, to disrupt the general flow of conversation (they should be nearing the end of their task at this point).

After 6 minutes have elapsed you should logout by clicking on the "Exit Chat" button.

**Now the computer will test you (3 times over) on the above information.**

**After this has ended, you will not be reminded of the information, so you must make an effort to learn the details provided by the software.**

Figure II-i: Participant Information Sheet (Part 1).

## Participant Information Sheet

### 2. The Disturbance

Now you will begin the disturbance. The software in front of you shows the gateway to the chat program. There is a single button in the centre of the page; the button is labelled "Enter Chat Room".

In a moment you will be asked to click on the button (at this point the chat room will be revealed). The chat program is very similar to ones that you may have used in the past. The main frame shows the chat "history" of all the posted messages. Underneath, is the chat entry box where you can type your own messages. Messages can be sent to the others by either clicking on the send button or by hitting the ENTER key.

Before you enter the chat room, you must LOGIN with the username and password. This username is added to any messages that you send to the chat room. You will be using the LOGIN details that you learnt earlier.

The first thing that you must do after you click the "Enter Chat Room" button, is to login using the EXACT same username and password that you have learnt. You may do this by clicking on the button labelled "LOGIN".

After you have logged in, the chat entry box will be enabled and you will be able to commit the disturbance. Remember there are 3 things that you will need to do.

Upon completion of your task you must click on the "Exit Chat" button and meet with the investigator who will be waiting outside.

If the software appears to break, or if you are unable to remember the username or password, then do not worry. Simply tell the investigator and the problem will be remedied.

**Now the investigator will show you a screenshot of the software that you will use and discuss any questions that you might have.**

Figure II-ii: Participant Information Sheet (Part 2).

## **Participant Information Sheet**

### **3. The Questionnaire**

Congratulations on completing the task! Now you will be presented with a piece of software that controls your navigation through a 3-Dimensional 'island'. Your movement can be controlled with the arrow keys on the keyboard.

The purpose of the exercise is to explore the island and locate a number of signs. There are five in total each placed at a specific location. For each sign, I would like you to write down on the answer paper exactly what is written on the sign. Additionally, I would like you to write down the co-ordinates where you discovered it (they can be approximate). The co-ordinates of where you are standing are displayed in the top left of the screen.

Throughout the island lives a certain breed of virtual cheetah. You must avoid these cheetahs at all cost because if they touch you, the game is over!

**You will be given 10 minutes for this exercise. Good Luck!**

Figure II-iii: Participant Information Sheet (Part 3).



## **Participant Information Sheet**

### **4. The Test**

Now you will be asked to learn some new phrases, six in total. In a moment, the computer will help you to remember them in a similar way to how it did at the beginning of the experiment.

Following this, you will be shown a piece of software that will list about 100 items, one after the other, in about one-second intervals.

If you see any of the 6 new phrases that you have learnt then you must click the LEFT MOUSE BUTTON

If you DO NOT see these phrases, then you must click the RIGHT MOUSE BUTTON

You must respond as quickly and as accurately as you can (under a second). If you spend longer than a second on each response, the message "Too Slow" will appear before the next item in the list.

You will repeat this Learn – Test procedure twice with different phrases each time.

**GOOD LUCK**

**You will now begin the learning procedure.**

Figure II-iv: Participant Information Sheet (Part 4).



Figure II-v: The Distracter Task.

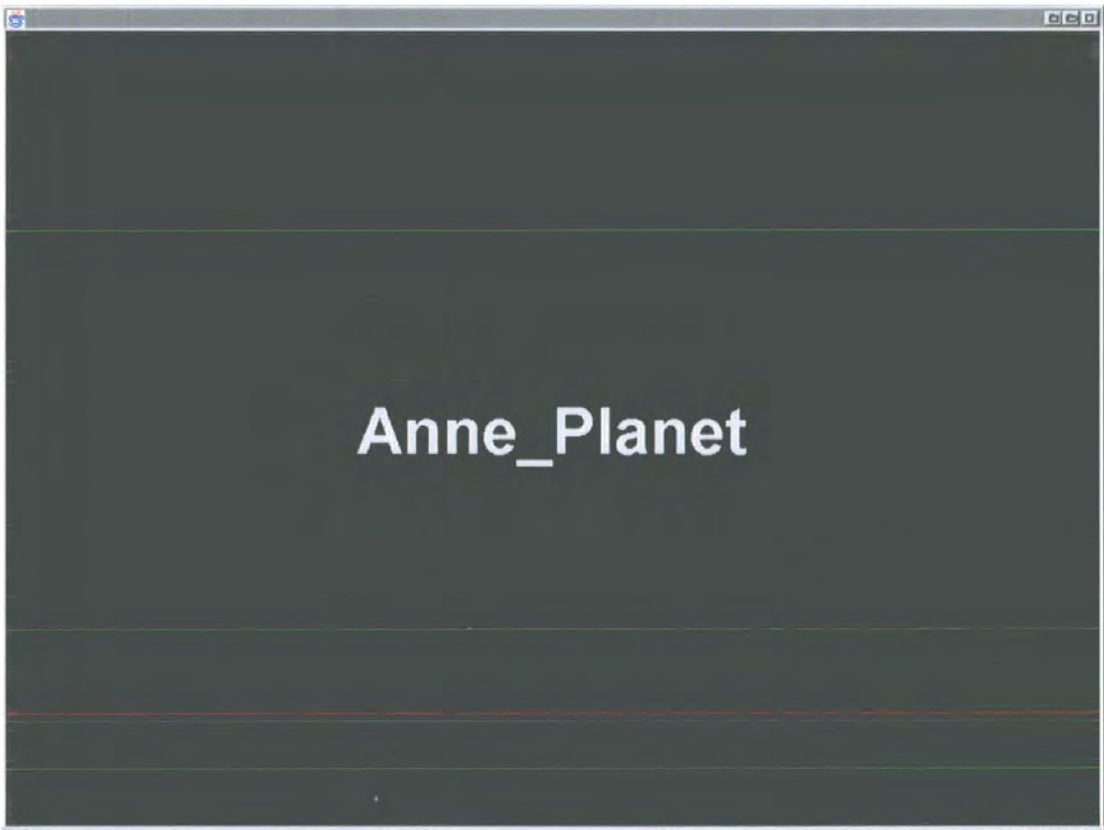


Figure II-vi: The Tester Software

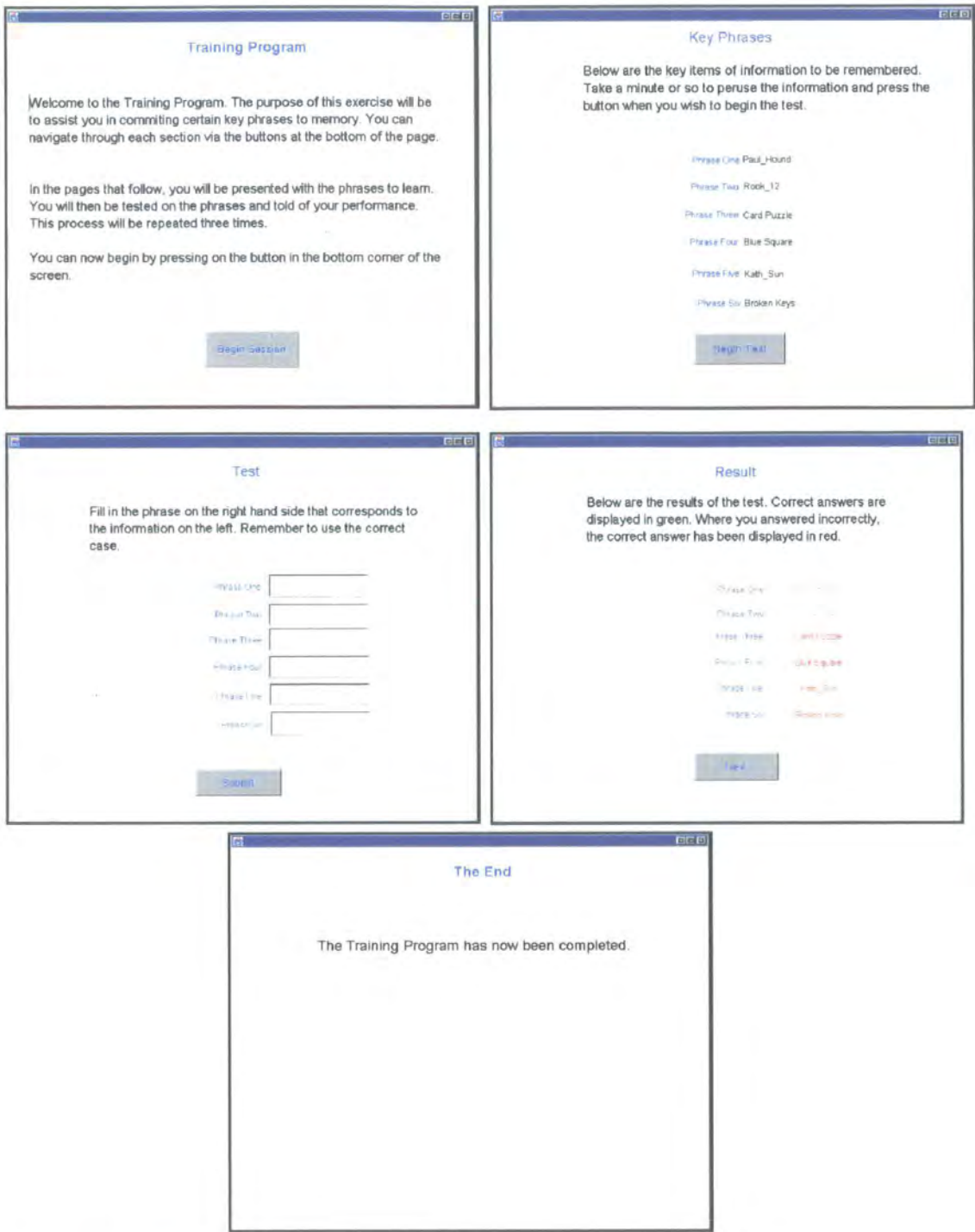


Figure II-vii: The Training Software.



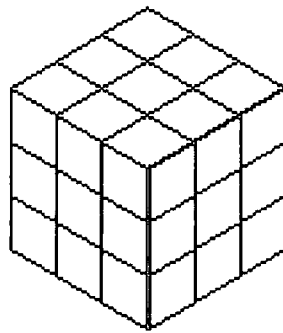
Figure II-viii: The Chat Room Simulator

Jane_Star	Lets get the cube together first :-)
Jane_Star	Think I better describe mine again.
Parson_S	My base is covered in yellow stickers rest is of mixed colours
Jonny_G	I think that stickers on the inside when put together match up. I also think that one face may be completely green (and one side red?????)
Jonny_G	Parson_S I think that that is significant
Sunny	Would Jane_Star' fit with Jonny_Gs to make a base?
Jane_Star	I've got a base of 5 cubes, with 3 cubes in a line, and 2 cubes next to them.
Jonny_G	what do others think about my idea of the colour stickers matching up?
Jane_Star	Mine will fit with So_phias to make the base, and then have one cube sticking up into the second layer in the corner.
Parson_S	could be rubix cube?
Sunny	could be, I've got one side of my three all red
Jonny_G	ah - Jane_Star that makes more sense
Sunny	the other sides are a mix of colours
So_phia	I have my one cube with all external sides yellow so it could be a corner bit
Jonny_G	ok - Sunny I ours may be together.
Parson_S	ok to describe the shape I have again a 2 * 2 *2 L
Jane_Star	Alternatively, mine could be the base with Jonny_Gs up the side?
Parson_S	Jonny the base of my shape is yellow so could that somehow fit with yours?
Sunny	mine is three cubes in a row
So_phia	so Sunny has a central row then ?
Jonny_G	Sunny have you got a green sticker that when slotted into my hole will produce a green sticker on the underside?
Jonny_G	Sunny is one of your ends green
Jonny_G	?
Jonny_G	So_phia - yes
Jane_Star	I don't think the pattern of dots is significant. Reading the question implies that each side is going to have a different pattern
Sunny	I've got two green stickers (one on each end) and a line
So_phia	so Jonny_G is top or bottom ?
Jonny_G	right - so that must slot into mine either way. This will mean that certainly one side is all green - the opposite side will be mixed / green
Jonny_G	lets say the bottom
Sunny	If we put mine and Jonny_Gs bit together we get a solid base and a column two high
Sunny	in the middle
Jonny_G	excellent
Jane_Star	Ok. If you put mine to as a side, then So_phias could be another side.
Jonny_G	yes
So_phia	yep I think so
Parson_S	mine has got to be a corner piece
Jane_Star	Getting hard to visualise this!
Jonny_G	Jane_Star is one of your sides all red?
Sunny	Does that all fit?
Jane_Star	All my sides have different colours
Jane_Star	i.e. more than one colour per side
Jonny_G	ok
So_phia	so if we have used mine, Jane_Star, Jonny_Gs and Sunnys that leaves Parson_Ss to slot on the top somehow
Jane_Star	Is anyone making a note of what it looks like so far?
Jonny_G	Jane_Star do any of your sides have the pattern blue-yellow-blue or blue-green-blue?
Parson_S	I think mine might slot over the column of two high
So_phia	I need pen and paper !
So_phia	I have Red Red Green on one of my 'long' sides
Parson_S	the column of two high being the shape produced by Sunny and Jonny_G's shapes
Jane_Star	Jonny_G-potentially. Ive got a row with a blue dot and a yellow dot.
So_phia	right I have pen & paper now shall we try again from the bottom ?
Jonny_G	Parson_S - I understand, so does it go round the column (rather than on top
Parson_S	If you imagine the L upside down then it could wrap round the column that way?
Jonny_G	So_phia - good idea. There is mine with Sunnys going though making a base of 3x3 with a column of 2 on top in the center

Figure II-ix: Sample of the Transcript used in the Simulation.

## Instructions

This is a group activity in which you will be asked to communicate with your team members solely through a computer-mediated chat program. When you are told to begin, you should open the box that has been presented to you. Inside you will find a piece of a puzzle. **YOU MUST NOT SHOW THIS PIECE TO ANY OF THE OTHER TEAM MEMBERS.** The solution to the puzzle is a 3 X 3 X 3 cube as shown below.



Using the chat program as your sole means of communication, you are to solve this puzzle with your other teammates. You will be given 30 minutes to:

1. Describe how each piece fits together in order to produce a cube. Your final solution should be entered into the chat program after the solution has been found, prefixed with the text "FINAL SOLUTION".
2. Each individual member must then describe the pattern of dots that fall on a particular face. No two team members are allowed to describe the same face. A team member can enter only one description, prefixed with the text "FACE"

**You will be paid £3 pounds for this activity. A bonus of £3 will be paid if all parts of the exercise have been completed for the team as a whole.**

Figure II-x: Instructions for the Block Puzzle Task.



# Appendix III

## Addendum for Chapter 7

Screenshots of the Chat Client used to conduct the interviews are presented in Figures III-i and III-ii. The Participant Information Sheet for this study has been provided in Figure III-iii.



Figure III-i: The Chat Client.

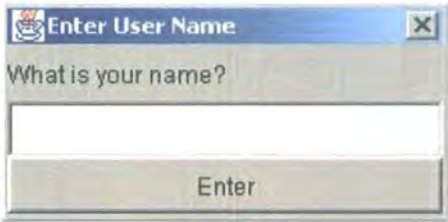


Figure III-ii: The Login Dialogue.

## Participant Information Sheet

Hello, thank you for volunteering to take part in the study! The purpose of this investigation is to learn how an interview might be carried out within a software-based "chat room". The study is being undertaken as part of a Ph.D. with the Department of Computer Science and if you have any queries upon its conclusion you should direct them to:

"Mr Simon J. Placks [s.j.placks@durham.ac.uk](mailto:s.j.placks@durham.ac.uk)"

Department of Computer Science

Science Site, University of Durham

Your participation is voluntary, and you are free to pull out at any time during or after the study without any need to provide a reason. If you do wish to pull out, then all record of the interview will be destroyed. The whole process should take about an hour and a half.

In a moment, you will be asked to watch a short clip from a motion picture. Try to remember as much as you can about the clip because this will be the topic of the interview. After you have watched the clip, the investigator (Simon) will introduce you to the chat software. After a brief familiarisation session, the interview will take place. In the interview, you will be asked to recall as much as you can from the video clip, but this is not a memory test and you will not be scored on the number of details remembered! The purpose of this interview is to assess the technique of interviewing through the software. During the interview, the investigator will be in an adjoining room. After the main interview, you will be reunited with the investigator and then informally interviewed about your perceptions of the process.

During the interview, you will be referred to by a pseudonym, generated automatically by the investigator. You should be careful not to disclose information that is personally identifying in any way. Similarly, the interview will solely concern the chat software and the film clip. The transcript of the interview will be recorded by the chat software and will be privately stored by the investigator until the research has been completed. A booklet that connects your real name to the pseudonym will also be kept, so that if you wish to pull out of the study later, the relevant data can be identified and destroyed.

This study is exploratory and I would be very interested to hear your opinions about the whole process at the end of the interview. If you have any questions please feel free to ask. In case of technical difficulty (i.e. the computer starts behaving strangely or crashes) then do not worry at all. The software is experimental and may occasionally break down for no apparent reason. Simply contact me in the following room. Similarly, I will contact you if for some reason I also have problems!

Finally, thanks again for your participation. If you have any further questions, please ask!

Figure III-iii: Participant Information Sheet.



Appendix IV

Addendum for Chapter 8

The statistics regarding deletion behaviours have been provided in Tables IV-i and IV-ii below. The Participant Information Sheet is provided in Figures IV-i and IV-ii.

<i>P</i>	<i>Condition</i>	<i>Total Keys</i>	<i>Total BKeys</i>	<i>Total Backspace Strings</i>	<i>Ratio<sub>b:total</sub> %</i>	<i>Ratio<sub>bstring:total</sub> %</i>	<i>Mean BString Length</i>	<i>S.D.</i>
1	Truthful	5026	437	233	8.69	4.64	1.88	1.62
	Fabricated	4776	265	164	5.55	3.43	1.62	0.97
2	Truthful	6982	449	174	6.43	2.49	2.58	3.16
	Fabricated	7677	641	232	8.35	3.02	2.76	3.14
3	Truthful	8009	409	102	5.11	1.27	4.01	5.47
	Fabricated	7447	556	113	7.47	1.52	4.92	7.89
4	Truthful	6447	670	252	10.39	3.91	2.66	3.13
	Fabricated	5980	615	235	10.28	3.93	2.62	3.17
5	Truthful	6299	1053	210	16.72	3.33	5.01	6.55
	Fabricated	5752	1175	204	20.43	3.55	5.76	7.83
6	Truthful	4355	391	106	8.98	2.43	3.69	6.62
	Fabricated	4292	466	120	10.86	2.80	3.88	8.11
7	Truthful	7723	347	206	4.49	2.67	1.68	0.88
	Fabricated	8156	314	184	3.85	2.26	1.71	1.08
8	Truthful	5563	154	66	2.77	1.19	2.33	2.56
	Fabricated	4868	102	44	2.10	0.90	2.32	2.15
9	Truthful	5925	779	248	13.15	4.19	3.14	3.40
	Fabricated	6440	741	293	11.51	4.55	2.53	2.95
10	Truthful	3738	249	71	6.66	1.90	3.51	4.76
	Fabricated	3894	229	70	5.88	1.80	3.27	5.27
11	Truthful	6510	674	173	10.35	2.66	3.90	7.87
	Fabricated	7106	493	216	6.94	3.04	2.28	2.38

Table IV-i: Deletion Statistics (Part 1).

<i>P</i>	<i>Condition</i>	<i>Total Keys</i>	<i>Total BKeys</i>	<i>Total Backspace Strings</i>	<i>Ratio<sub>b:total</sub> %</i>	<i>Ratio<sub>bstring:total</sub> %</i>	<i>Mean BString Length</i>	<i>S.D.</i>
12	Truthful	4152	280	127	6.74	3.06	2.20	2.12
	Fabricated	6351	372	194	5.86	3.05	1.92	1.87
13	Truthful	6086	584	201	9.60	3.30	2.91	5.44
	Fabricated	4932	596	147	12.08	2.98	4.05	9.95
14	Truthful	5373	456	129	8.49	2.40	3.53	6.64
	Fabricated	5195	586	142	11.28	2.73	4.13	9.22
15	Truthful	5458	566	96	10.37	1.76	5.90	9.09
	Fabricated	5231	776	102	14.83	1.95	7.61	10.84
16	Truthful	8239	543	143	6.59	1.74	3.80	5.36
	Fabricated	7249	461	116	6.36	1.60	3.97	5.22
17	Truthful	4975	388	140	7.80	2.81	2.77	5.44
	Fabricated	4126	203	77	4.92	1.87	2.64	3.06
18	Truthful	5340	177	66	3.31	1.24	2.68	2.64
	Fabricated	5720	398	103	6.96	1.80	3.86	5.43
19	Truthful	5761	475	208	8.25	3.61	2.28	1.91
	Fabricated	6404	569	250	8.89	3.90	2.28	2.11
20	Truthful	4642	249	94	5.36	2.02	2.38	1.98
	Fabricated	4361	376	115	8.62	2.64	3.27	4.48
21	Truthful	6147	549	99	8.93	1.61	5.43	12.12
	Fabricated	4828	284	102	5.88	2.11	2.78	3.83
22	Truthful	4415	639	145	14.47	3.28	4.10	7.92
	Fabricated	4460	379	114	8.50	2.56	3.32	4.99
23	Truthful	3575	333	121	9.31	3.38	2.75	3.59
	Fabricated	3871	476	127	12.30	3.28	3.75	8.52
24	Truthful	8099	496	169	6.12	2.09	2.93	3.33
	Fabricated	5896	300	113	5.09	1.92	2.65	2.65

Table IV-ii: Deletion Statistics (Part 2).

## Participant Information Sheet

Hello! Thank you for volunteering to take part in the study. The purpose of this study is to investigate the ways in which people talk about fictitious events when using chat room software. Specifically, the study is looking at how the telling of a fictitious event can differ from the telling of an actually experienced event. The whole process should take about an hour and a half.

In a moment, you will be interviewed through a computer chat program. During the interview, I will ask you about two day trips that you have been on, one after the other. What I would like you to do, is to completely fabricate one of these accounts. The other account should be related as accurately as you remember. It is up to you to decide whether to tell the true story first or second, but most importantly do not reveal to me (at any point) which story was true, and which was fabricated (this question will be asked by the software system). You will be provided with an equal amount of preparation time for both stories. The stories should be picked to avoid any experiences that you found emotionally distressing and should refer to at least one other person.

The Interview will be conducted through a chat room program, probably similar to ones that you have used in the past. The chat room program is a private system, which can only be accessed by you and me. This is an unusual scenario, and it is possible that the software will fail on occasion. If this does occur then do not worry, the interviewer will be able to easily restart the system.

For the "made up" day trip, I would like you to fabricate a significant part of the story (so for example don't simply fabricate dates and times). In this respect, the fabrication is similar to that used by actors and actresses when performing. It is essential that you try to make this account as convincing as possible.

The chat room program we will be using has a number of features unique to this study. During the session, the program will be recording information about the interview, such as the transcript as well as other data. At the end of the interview, NONE of this data will be saved automatically. Instead, the software will ask which data should be saved to disk. At this point, you will be asked again which data you permit to be stored. This decision is entirely up to you. The interviewer, will strip out any personal information from the transcript before it is used in analysis (for example, all names will be substituted).

Figure IV-i: Participant Information Sheet (Part 1).

## Participant Information Sheet

cont...

The role of volunteers like you in this study is vital; it could not take place otherwise. The study is being conducted as part of a body of Ph.D. research at the Computer Science Department of Durham University. My contact details are provided below. The study is entirely voluntary, and you are free to pull out at any time. If at any time you decide to withdraw, then simply contact me at my address below, and you will be sent confirmation upon deletion of the data.

My contact details are:

Mr Simon J. Placks  
Postgraduate  
Department of Computer Science  
Science Site  
University of Durham,  
Email: s.j.placks@dur.ac.uk

Once again, thank you for your help! If you have any more questions about the study then feel free to ask!

Figure IV-ii: Participant Information Sheet (Part 2).

## Glossary

---

<b><i>A/S/L</i></b>	Commonly used acronym within CMC to request the age, sex, and location of the collocutor.
<b><i>Affordance</i></b>	A term coined by Gibson (1979) to refer to an opportunity for action that a particular artefact represents for a particular organism.
<b><i>AJE</i></b>	Automated Joining Event. A term coined by Rintel et al. (2001) to refer to the announcement made by a chat system when a new user joins the room.
<b><i>BBS</i></b>	Bulletin Board Service
<b><i>CBBS</i></b>	Computerized Bulletin Board Service. One of the first BBS systems.
<b><i>CBCA</i></b>	Criteria-based Content Analysis. The content analysis technique that used with transcripts generated as part of a Statement Validity Assessment procedure.
<b><i>CEP</i></b>	Channel Entry Phase. A term coined by Rintel et al. (2001) to refer to the opening stages of discursive interaction following an Automated Joining Event.
<b><i>CI</i></b>	Cognitive Interview. A semi-structured investigative interviewing technique developed by Fisher and Geiselman (1992).
<b><i>CMC</i></b>	Computer-Mediated Communication. Any form of computer-based communication that facilitates human interpersonal discourse. Within this thesis, the term CMC is used to refer to text-based CMC systems.
<b><i>Cognitive Load</i></b>	The level of cognitive activity that is being undertaken by the brain.

<b><i>Context Interactivity</i></b>	A term used within Interpersonal Deception Theory to describe the inherent level of interactivity of a communicative socio-technical context.
<b><i>Context Reinstatement</i></b>	A stage of the Cognitive Interview whereby the interviewee is asked to mentally reflect upon a particular remembered event.
<b><i>Continuer</i></b>	A term used in Conversation Analysis to describe a short verbal device that prompts a speaker to continue their discourse.
<b><i>Controlled Lie-Detection</i></b>	Lie-detection that requires careful control of the manner in which a potential deceiver would need to lie.
<b><i>CQT</i></b>	Control Questions Test. A polygraph test developed by Reid and Inbau (1977).
<b><i>Duping Delight</i></b>	Term coined by Ekman (2001) to describe the experience of exhilaration that accompanies a successful deceit.
<b><i>ECI</i></b>	Enhanced Cognitive Interview. An improved form of the Cognitive Interview that provides a precisely defined interviewing schedule.
<b><i>EEG</i></b>	Electroencephalograph. A device used to record Event Related Potentials within the brain.
<b><i>EIES</i></b>	Electronic Information Exchange System. An early CMC system.
<b><i>Emoticons</i></b>	Ideographs that represent emotional reactions and that are depicted using textual characters.
<b><i>ERP</i></b>	Event-related Potential. A specific waveform recorded by an EEG in response to a specific stimulus or 'event'.
<b><i>GKT</i></b>	Guilty Knowledge Test. A polygraph test developed by Lykken (1959,1960).
<b><i>Ground Truth</i></b>	The objective reality (in a realist sense) that defines a state of the world that is considered to be the truth, and that can be compared against a particular statement in order to test its consistency with that truth.

<b><i>Hyperpersonal</i></b>	A term coined by Walther (1996) to describe the inflated levels of intimate interpersonal interaction that can occur within CMC contexts.
<b><i>ICFS</i></b>	Independent Case Fact Scales. A term coined by Lamb et al. (1997) to describe a portfolio of evidence that could represent the ground truth within field-based lie-detection studies.
<b><i>ICQ</i></b>	A modern CMC system that facilitates quasi-synchronous, synchronous, and asynchronous discourse.
<b><i>IDT</i></b>	Interpersonal Deception Theory. A theory developed by Buller and Burgoon (1996) that describes interpersonal deceptive encounters.
<b><i>Interactivity Principle</i></b>	A principle defined within Interpersonal Deception Theory that describes the relationship between the level of interactivity afforded by a medium, and the subsequent levels of suspicion and deceptive success perceived within a deceptive encounter.
<b><i>IRC</i></b>	Internet Relay Chat. A popular text-based CMC system.
<b><i>Latency Effect</i></b>	Slowed reaction times during a task.
<b><i>Leakage</i></b>	The clues to deception that escape the suppressive control of the deceiver.
<b><i>Lexical Diversity</i></b>	The range of vocabulary used by a speaker. Lexical diversity is often assessed by calculating the type-token ratio.
<b><i>LSI</i></b>	Laboratory for Scientific Investigations. The commercial company that developed SCAN.
<b><i>Mock Crime</i></b>	An experimental procedure in which a participant is engaged in committing a make-believe crime that can later serve as the subject for a lie-detection test.
<b><i>MOO</i></b>	MUD Object-Oriented. An extension to the basic MUD framework that allows a greater power in creating and manipulating objects within the text-based universe.

<b><i>MUD</i></b>	Multi-User Domain. A virtual environment that allows users to interact with one another within a world defined through text.
<b><i>MUD Wizard</i></b>	A MUD user who is afforded special privileges, usually relating to system administration, member support, and disciplinary functions.
<b><i>Naturalistic Lie-Detection</i></b>	Lie-detection whereby the technique is adapted so that can be robust enough to detect a variety of lies without directly manipulating the deceptive context.
<b><i>Newsmark</i></b>	A term used within Conversation Analysis to refer to a short verbal device that informs the speaker that the listener has understood their previous utterance.
<b><i>OR</i></b>	Orienting Reflex. The reflex response that is triggered upon elicitation of a novel stimulus.
<b><i>Phantom Adjacency Pairs</i></b>	A pair of utterances, one from each collocutor, that seem to directly refer to one another but that were not intended to be connected in this manner by either speaker. Such occurrences are usually found during misalignment.
<b><i>Pseudonymity</i></b>	Communication by which each speaker is referred to solely by pseudonym.
<b><i>QS-CMC</i></b>	Quasi-synchronous CMC. A term coined by Garcia and Jacobs (1999) to describe text-based CMC that allows a speaker to edit their utterance before it is posted to a listener during a real-time conversation.
<b><i>RM</i></b>	Reality Monitoring. A content analysis technique that assesses veridicality.
<b><i>RT</i></b>	Response Time. The time taken for a participant to click a button following the elicitation of a stimulus.
<b><i>SCAN</i></b>	Scientific Content Analysis. A commercial lie-detection technique developed by LSI.



<b><i>SIDE</i></b>	Social Identification and Deindividuation Effects. A model developed by Spears and Lea (1992) to describe the relationship between identity salience and deindividuation effects.
<b><i>SIP</i></b>	Social Information Processing. A theoretical approach to CMC that was presented by Walther (1992) and that emphasised the adaptive abilities of collocutors to attain social and relational information through communication media.
<b><i>SRA</i></b>	Statement Reality Analysis. An early form of Statement Validity Assessment.
<b><i>SVA</i></b>	Statement Validity Assessment. A three-part truth-validation technique consisting of an investigative interview, content analysis, and validity checklist.
<b><i>TRP</i></b>	Transition-Relevance Place. A term used within Conversation Analysis to describe the period in which participants in a conversation perceive an opportunity for turn transition.
<b><i>Type-token Ratio</i></b>	The ratio of the total number of words to the total number of unique words.
<b><i>Undeutsch Hypothesis</i></b>	The hypothesis coined by Udo Undeutsch that predicts quantitative and qualitative differences between statements produced by deceivers and truth-tellers.
<b><i>VCL</i></b>	Validity Checklist. The part of SVA that assess the level of certainty that can be ascribed to CBCA scores.

## References

---

[BBC 24/10/2000] Warning as Net Predator Sentenced. Published online at

<http://news.bbc.co.uk/1/hi/uk/987418.stm> *last accessed 19/09/03*

[BBC 15/03/2001] Internet Chat Danger for Children. Published online at

<http://news.bbc.co.uk/1/hi/uk/1222257.stm> *last accessed 19/09/03*

[CNN 24/07/2000] Police Hook Pedophile on Web in Five Seconds. By Russel Lissau, Daily Herald. Published online at

<http://www.cnn.com/2000/LOCAL/westcentral/07/24/ahd.police.web/index.html> *last accessed 19/09/03*

[Home Office, 1998] Sex Offending Against Children: Understanding the Risk. Police Research Series Paper 99. By Don Grubin, Policing and Reducing Crime Unit (Home Office, Research Development Statistics). Published online at

<http://www.homeoffice.gov.uk/rds/prgpdfs/fprs99.pdf> *last accessed 19/09/03*

[ZDNET 20/11/2002] Paedophiles Given Notice to Quit Chatrooms. By Eugene Lacey.

Published online at <http://comment.zdnet.co.uk/eugenelacey/0,39020763,2126231,00.htm> *last accessed 19/09/03*

---

Akehurst, L., Kohnken, G. and Hofer, E. (2001) Content Credibility of Accounts Derived from Live and Video Presentations. Legal and Criminological Psychology (6), 65-83.

Alonso-Quecuty, M.L (1992) Deception Detection and Reality Monitoring: A New Answer to an Old Question? In F. Losel, D. Bender, and T. Bliesener (eds.) Psychology and Law: International Perspectives. Berlin, de Gruyter.

Alonso-Quecuty, M.L. (1996) Detecting Fact From Fallacy in Child and Adult Witness Accounts. In G. Davies, S. Lloyd-Bostock, M. McMurran, and C. Wilson (eds.) Psychology, Law, and Criminal Justice: International Developments in Research and Practice. Berlin, de Gruyter.

Argyle K. and Shields, R. (1996) Is There a Body in the Net? In R. Shields (ed.) Cultures of Internet: Virtual Spaces, Real Histories, Living Bodies. London, Sage.

- Aronsen E., Wilson, T.D. and Brewer, M.B. (1999) Experimentation in Social Psychology. In D.T. Gilbert, S.T. Fiske, and G. Lindzey (eds.) The Handbook of Social Psychology. 4<sup>th</sup> ed. New York, Oxford University Press.
- Anson, D.A., Golding, S.L. and Gully, K.J. (1993) Law and Human Behavior (17) no.3, 331-341.
- Bailey, F.G. (1991) The Prevalence of Deceit. 1<sup>st</sup> ed. New York, Cornell University Press.
- Baldwin, J. (1993) Police Interview Techniques: Establishing Truth or Proof? British Journal of Criminology (33), 325-352.
- Barland, G.H. and Raskin, D.C. (1975) An Evaluation of Field Techniques in Detection of Deception. Psychophysiology (12) no. 3, 321-330
- Barnes, J.A. (1994) A Pack of Lies: Towards a Sociology of Lying. 1<sup>st</sup> ed. Cambridge, Cambridge University Press.
- Baron, N.S. (1984) Computer-Mediated Communication as a Force in Language Change. Visible Language (18) no. 2. Spring, 118-141
- Baumeister, R.F. (1998) The Self. In D.T. Gilbert, S.T. Fiske, and G. Lindzey (eds.) The Handbook of Social Psychology. 4<sup>th</sup> ed. New York, Oxford University Press.
- Baym, N. (1995) The Emergence of Community in Computer-Mediated Communication. In S. Jones (ed.) Cybersociety: Computer Mediated Communication and Community. London, Sage.
- Ben-Shakhar, G. (2002) A Critical Review of the Control Questions Test (CQT). In: M. Kleiner (ed.) Handbook of Polygraph Testing. 1<sup>st</sup> ed. London, Academic Press.
- Ben-Shakhar, G. and Elaad, E. (2002) The Guilty Knowledge Test (GKT) as an Application of Psychophysiology: Future Prospects and Obstacles. In: M. Kleiner (ed.) Handbook of Polygraph Testing. 1<sup>st</sup> ed. London, Academic Press.
- Ben-Shakhar, G. and Furedy, J.J. (1990) Theories and Applications in the Detection of Deception: A Psychophysiological and International Perspective. 1<sup>st</sup> ed. New York, Springer Press.
- Black, M. (1983) The Prevalence of Humbug and Other Essays. Ithaca, Cornell University Press.

Blackman, R. (1966) The Effect of the Orienting Reaction on Disjunctive Reaction Time. Psychonomic Science (4), 411-412.

Bok, S. (1999) Lying: Moral Choice in Public and Private Life. 3<sup>rd</sup> ed. New York, Vintage Books.

Bowker, N. and Tuffin, K. (2002) Disability Discourses for Online Identities. Disability and Society (17) no.3, 327-344.

Boychuk, T.D. (1991) Criteria-based Content Analysis of Children's Statements About Sexual Abuse: A Field Based Evaluation Study. Unpublished doctoral dissertation, Arizona State University.

Brewer, M.B. and Brown, R.J. (1998) Intergroup Relations. In D.T. Gilbert, S.T. Fiske, and G. Lindzey (eds.) The Handbook of Social Psychology. 4<sup>th</sup> ed. New York, Oxford University Press.

Bromberg, H. (1996) Are MUDs Communities? Identity, Belonging and Consciousness in Virtual Worlds. In R. Shields (ed.) Cultures of Internet: Virtual Spaces, Real Histories, Living Bodies. 1<sup>st</sup> ed. London, Sage.

Buller, D.B. and Burgoon, J.K. (1996) Interpersonal Deception Theory. Communication Theory (6) no.3, 203-242.

Buller D.B., Burgoon, J.K., Buslig, A. and Roiger, J. (1996) Testing Interpersonal Deception Theory: The Language of Interpersonal Deception. Communication Theory (6) no.3, 268-288.

Burgoon, J.K. and Buller D.B. (1996) Reflections on the Nature of Theory Building and the Theoretical Status of Interpersonal Deception Theory. Communication Theory (6) no.3, 310-328.

Burgoon, J.K., Stern, L.A., and Dillman, L. (1995) Interpersonal Adaptation: Dyadic Patterns of Interaction. New York, Cambridge University Press.

Burgoon, J.K., Buller, D.B., Ebesu, A.S., White, C.H. and Rockwell, P.A. (1996a) Testing Interpersonal Deception Theory. Effects of Suspicion on Communication Behaviors and Perceptions. Communication Theory (6) no.3, 243-267.

Burgoon, J.K., Buller, D.B., Floyd, K. and Grandpre, J. (1996b) Deceptive Realities: Sender, Receiver, and Observer Perspectives in Deceptive Conversations. Communication Research (23) no.6, 724-748.

- Burgoon, J.K., Buller, D.B., White, C.H., Afifi, W., and Buslig, A.L.S. (1999) The Role of Conversational Involvement in Deceptive Interpersonal Interactions. Personality and Social Psychology Bulletin (20) no. 6. June, 669-685.
- Burgoon, J.K., Stoner, G.M., Bonito, J.A., and Dunbar, N.E. (2003) Trust and Deception in Mediated Communication. Proceedings of the Hawaii International Conference on Systems Science, Kona, HI.
- Campbell, D.T. (1967) Stereotypes and Perception of Group Differences. American Psychologist (22), 817-829.
- Carpenter, R.H. (1990) The Statistical Profile of Language Behavior with Machiavellian Intent or While Experiencing Caution and Avoiding Self-Incrimination. Annals of the New York Academy of Sciences (606), 5-17.
- Carver C.S. and Scheier, M.F. (1981) Attention and Self-Regulation: A Control Theory Approach to Human Behavior. New York, Springer-Verlag.
- Chen, P. and Hinton, S.M. (1999) Realtime Interviewing Using the World Wide Web. Sociological Research Online (4) no.3.
- Columbia Pictures (1962) Lawrence of Arabia. Produced by Sam Spiegel. Directed by David Lean. © Columbia Pictures.
- Cornwell, B. and Lundgren, D.C. (2001) Love on the Internet: Involvement and Misrepresentation in Romantic Relationships in Cyberspace vs. Realspace. Computers in Human Behavior (17), 197-211.
- Curtis, P. (1992) MUDing: Social Phenomena in Text-based Virtual Realities. In P. Ludlow (ed.) High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace. Cambridge, MIT Press.
- Davis, R.C. (1961) Physiological Responses as a Means of Evaluating Information. In A.D. Biderman and H. Zimmer (eds.) The Manipulation of Human Behavior. New York, Wiley.
- DePaulo, B.M., Kashy, D.A., Kirkendol, S.E., Wyer, M.M., and Epstein, J.A. (1996a) Lying in Everyday Life. Journal of Personality and Social Psychology (70), 979-995.

DePaulo B.M., Ansfield, M.E., and Bell, K.L. (1996b) Theories About Deception and Paradigms for Studying It: A Critical Appraisal of Buller and Burgoon's Interpersonal Deception Theory and Research. Communication Theory (6) no.3, 296-309.

Dibble, J. (1996) A Rape in Cyberspace; or How an Evil Clown, a Haitian Trickster Spirit, Two Wizards, and a Cast of Dozens Turned a Database into a Society. In. P. Ludlow (ed.) High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace. 1<sup>st</sup> ed. Cambridge, The MIT Press.

Diener, E. (1979) Deindividuation, Self-Awareness and Disinhibition. Journal of Personality and Social Psychology (37), 1160-1171.

Diener, E. (1980) De-individuation: The Absence of Self-Awareness and Self-Regulation in Group Members. In P. Paulus (ed.) The Psychology of Group Influence. Hillsdale, Erlbaum.

Donath, J.S. (2001) Identity of Deception in the Virtual Community. In P. Kollock and M. Smith (eds) Communities in Cyberspace. London, Routledge.

Douglas, K.M. and McGarty, C. (2001) Identifiability and Self-Presentation: Computer-Mediated Communication and Intergroup Interaction. British Journal of Social Psychology (40), 399-416.

Douglas, J.D. (1976) Investigative Social Research. Beverly Hills, Sage.

Ekman, P. (2001) Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage. 3<sup>rd</sup> ed. New York, W.W.Norton & Company, Inc.

Elaad, E. (1998) The Challenge of the Concealed Knowledge Polygraph Test. Expert Evidence (6), 161-187.

Farwell, L.A. and Donchin, E. (1991) The Truth Will Out: Interrogative Polygraphy ("Lie Detection") with Event-Related Brain Potentials. Psychophysiology (28) no. 5, 531-547.

Ferguson, J.A. (1977) PLANET: A Computer Conferencing System and its Evaluation through a Case Study. Behavior Research Methods and Instrumentation (9) no.2, 92-95

Fisher, R.P. and Geiselman, R.E. (1992) Memory-Enhancing Techniques for Investigative Interviewing: The Cognitive Interview. 1<sup>st</sup> ed. Springfield, Charles C Thomas.

- Fisher, R.P., Geiselman, R.E. and Raymond, D.S. (1987) Critical Analysis of Police Interview Techniques. Journal of Police Science and Administration (15), 291-297.
- Franco, V., Piirto, R., Hu, H., and Lewenstein, B.V. (1995) Anatomy of a Flame: Conflict and Community Building on the Internet. IEEE Technology and Society Magazine Summer, 12-21
- Friedman, H.S., Riggio, R.E. and Casella, D.F. (1988) Nonverbal Skill, Personal Charisma, and Initial Attraction. Personality and Social Psychology Bulletin (14), 203-211.
- Garcia, A.C. and Jacobs, J.B. (1999) The Eyes of the Beholder: Understanding the Turn-Taking System in Quasi-Synchronous Computer-Mediated Communication. Research on Language and Social Interaction (32) no.4, 337-367
- Gergen, K. (1991) The Saturated Self: Dilemmas of Identity in Contemporary Life. New York, Basic.
- Gibson, J.J. (1979) The Ecological Approach to Visual Perception. 1<sup>st</sup> ed. New Jersey, Lawrence Erlbaum Associates Inc.
- Gibson, W. (1984) Neuromancer. London, Victor Gollancz Ltd.
- Ginton, A., Daie, N., Elaad, E., and Ben-Shakhar, G. (1982) A Method for Evaluating the Use of the Polygraph in a Real-Life Situation. Journal of Applied Psychology (67) no.2, 131-137.
- Goffman, E. (1975) Frame Analysis: An Essay on the Organisation of Experience. Harmondsworth, Penguin.
- Greenberg, J. and Pyszczynski, J. (1986) Persistent High Self-Focus After Failure and Low Self-Focus After Success: The Depressive Self-Focusing Style. Journal of Personality and Social Psychology (50), 1039-1044.
- Grint, K. and Woolgar, S. (1997) The Machine at Work. 1<sup>st</sup> ed. Cambridge, Polity Press.
- Hall, G.S. (1891) Children's Lies. Pedagogical Seminary (1), 211-218.
- Haraway, D. (1991) The Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. In Simians, Cyborgs, and Women. New York, Routledge.

- Hardey, M. (2002) Life Beyond the Screen: Embodiment and Identity Through the Internet. The Sociological Review. 570-585
- Hartung, J. (1988) Deceiving Down: Conjectures on the Management of Subordinate Status. In J.S. Lockard and D.L. Paulhus (eds.) Self-Deception: An Adaptive Mechanism? Englewood Cliffs, Prentice-Hall.
- Herring, S.C. (1999) Interactional Coherence in CMC. Journal of Computer-Mediated Communication (4), no.4.
- Hiemstra, G. (1982) Teleconferencing, Concern for Face, and Organizational Culture. In M. Burgoon (ed.) Communication Yearbook 6. Beverly Hills, Sage.
- Hiltz, S.R. and Turroff, M. (1978) The Network Nation. 1<sup>st</sup> ed. Addison-Wesley Publishing Company Inc.
- Hiltz, S.R. (1978) The Human Element in Computerized Conferencing Systems. Computer Networks (2), 421-428.
- Höfer, E., Akehurst, L., and Metzger, G. (1996) Reality Monitoring: A Chance For Futher Development of CBCA? Paper presented at the Annual Meeting of the European Association on Psychology and Law. August. Siena, Italy,
- Hollien, H. (1990) The Acoustics of Crime. The New Science of Forensic Phonetics. 1<sup>st</sup> ed. New York, Plenum Press.
- Horowitz, S.W., Lamb, M.E., Esplin, P.W., Boychuk, T.D., Krispin O. and Reiter-Lavery, L. (1997) Legal and Criminological Psychology (2), 11-21.
- Horowitz, S.W. (1998) Reliability of Criteria-based Content Analysis of Child Witness Statements: Response to Tully. Legal and Criminological Psychology (3), 189-191.
- Horvath, F.S. and Reid, J.E. (1971) The Reliability of Polygraph Examiner Diagnosis of Truth and Deception. The Journal of Criminal Law, Criminology and Police Science (62) no. 2, 276-281
- Horvath, F.S. (1977) The Effect of Selected Variables on Interpretation of Polygraph Records. Journal of Applied Psychology (62) no. 2, 127-136



- Hutchby, I. (2001) Conversation and Technology: From the Telephone to the Internet. 1<sup>st</sup> ed. Cambridge, Polity Press.
- Inbau, F.E., Reid J.E., and Buckley, J.P. (1986) Criminal Interrogation and Confessions. Baltimore, Williams & Wilkins.
- Ingram, R.E. (1990) Self-Focused Attention in Clinical Disorders: Review and a Conceptual Model. Psychological Bulletin (107), 156-176.
- Ito, M. (1997) Virtually Embodied: The Reality of Fantasy in a Multi-User Dungeon. In D. Porter (ed.) Internet Culture. Seattle, Seal Press.
- Jaffe, J.M., Lee, Y., Huang, L., and Oshagan, H. (1999) Gender Identification, Interdependence, and Pseudonyms in CMC: Language Patterns in an Electronic Conference. The Information Society (15), 221-234
- Johansen, R., Vallee, J. and Spangler, K. (1979) Electronic Meetings: Technical Alternatives and Social Choices. 1<sup>st</sup> ed. London, Addison-Wesley.
- Johnson, M.K. and Raye, C.L. (1981) Reality Monitoring. Psychological Review (88) no. 1, 67-85.
- Johnson, M.K. and Raye, C.L. (1998) False Memories and Confabulation. Trends in Cognitive Sciences (2), 137-145.
- Johnson, M.K., Hashtroudi, S. and Lindsay, D.S. (1993) Source Monitoring. Psychological Bulletin (114), 3-28.
- Jones, E.E. and Gerard, H.B. (1967) Foundations of Social Psychology. 1<sup>st</sup> ed. New York, Wiley.
- Kiesler, S., Siegel, J., and McGuire, T.W. (1984) Social Psychological Aspects of Computer-Mediated Communication. American Psychologist (39) no. 10, 1123-1134
- Knowles, P.L. and Smith, D.L. (1982) The Ecological Perspective Applied to Social Perception: Revision of a Working Paper. Journal for the Theory of Social Behaviour (12) no.1, 53-78.

- Köhnken, G., Schimossek, E., Aschermann, E., and Höfer, E. (1995) The Cognitive Interview and the Assessment of the Credibility of Adults' Statements. Journal of Applied Psychology (80) no.6, 671-684.
- Lamb M.E., Sternberg, K.J., Esplin, P.W., Hershkowitz, I., Orbach, Y., and Hovav, M. (1997) Criterion-Based Content Analysis: A Field Validation Study. Child Abuse and Neglect (21) no.3, 255-264.
- Landry, K.L. and Brigham, J.C. (1992) The Effect of Training in Criteria-Based Content Analysis on the Ability to Detect Deception in Adults. Law and Human Behavior (16) no.6, 663-676.
- Lea, M. and Spears, R. (1991) Computer-Mediated Communication, De-individuation and Group Decision-Making. International Journal of Man-Machine Studies (34), 283-301.
- Lea, M., Spears, R., and de Groot, D. (2001) Knowing Me, Knowing You: Anonymity Effects on Social Identity Processes Within Groups. Personality and Social Psychology Bulletin (27) no.5, 526-537.
- Liu, Y., Ginther, D., and Zelhart, P. (2001) How Do Frequency of Messaging Affect Impression Development in Computer-Mediated Communication? Journal of Universal Computer Science (7) no. 10, 893-913.
- Lykken, D.T. (1959) The GSR in the Detection of Guilt. Journal of Applied Psychology (43), 385-388.
- Lykken, D.T. (1960) The Validity of the Guilty Knowledge Technique: The Effects of Faking. Journal of Applied Psychology (44), 258-262.
- Lykken, D.T. (1974) Psychology and the Lie-Detection Industry. American Psychologist (29), 725-739.
- Lykken, D.T. (1979) The Detection of Deception. Psychological Bulletin (86) no.1, 47-53.
- Lynn, R. (1966) Attention, Arousal and the Orientation Reaction. 1<sup>st</sup> ed. Oxford, Pergamon Press.
- Mann, C. and Stewart, F. (2000) Internet Communication and Qualitative Research: A Handbook for Researching Online. 1<sup>st</sup> ed. London, Sage.

- Markham, A.N. (1998) Life Online: Researching Real Experience in Virtual Space. 1<sup>st</sup> ed. Walnut Creek, AltaMira Press.
- Matheson, K. and Zanna, M.P. (1988) The Impact of Computer-Mediated Communication on Self-Awareness. Computers in Human Behaviour (4), 221-233
- McArthur, L.Z. and Baron, R.M. (1983) Toward an Ecological Theory of Social Perception. Psychological Review (90), 215-247.
- McRae, S. (1997) Flesh Made Word: Sex, Text, and the Virtual Body. In D. Porter (ed.) Internet Culture. Seattle, Seal Press.
- McGrath, J.E. (1991) Time, Interaction, and Performance (TIP): A Theory of Groups. Small Group Research (22), 147-174.
- McGuire, T.W., Kiesler, S., and Siegel, J. (1987) Group and Computer-Mediated Discussion Effects in Risk Decision Making. Journal of Personality and Social Psychology (52) no.5, 917-930.
- McKenna, K.Y.A, and Bargh, J.A. (1998) Coming Out in the Age of the Internet: Identity "Demarginalization" Through Virtual Group Participation. Journal of Personality and Social Psychology (75), no. 3, 681-694.
- Memon, A. and Higham, P. (1999) A Review of the Cognitive Interview. Psychology, Crime and Law, (5), 177-196.
- Michaelis-Arntzen, E. (1987) Unglaubwuerdige Zeugenaussagen. Forensia, 8, 73-80.
- Milne, R. and Bull, R (1999) Investigative Interviewing: Psychology and Practice. 1<sup>st</sup> ed. Chichester, John Wiley & Sons Ltd.
- Mitchell K.J. and Johnson, M.K. (2000) Source Monitoring: Attributing Mental Experiences. In E. Tulving and F.I.M. Craik (eds.) The Oxford Handbook of Memory. Oxford, Oxford University Press.
- Morley, I.E. and Stephenson, G.M. (1969) Interpersonal and Interparty Exchange: A Laboratory Situation of an Industrial Negotiation at the Plant Level. British Journal of Psychology (6), 543-545.

- Moschovitz, C.J.P., Poole, H., Schuyler, T., and Senft, T.M. (1999) History of the Internet: A Chronology, 1843 to the Present. 1<sup>st</sup> ed. Santa Barbara, ABC-CLIO.
- Myers, D. (1987) "Anonymity is Part of the Magic": Individual Manipulation of Computer-Mediated Communication Contexts. Qualitative Sociology (10) no.3. Fall, 251-266
- Norman, D.A. (1988) The Psychology of Everyday Things. New York, Basic Books.
- O'Connor, H. and Madge, C. (2001) Cyber-Mothers: Online Synchronous Interviewing Using Conferencing Software. Sociological Research Online (5) no.4.
- van Olst, E.H. (1971) The Orienting Reflex. The Hague, Mouton.
- Orwell, G. (1984) 1984. Oxford, Clarendon Press.
- Parker, A.D. and Brown, J. (2000) Detection of Deception: Statement Validity Analysis as a means of Determining Truthfulness or Falsity of Rape Allegations. Legal and Criminological Psychology (5), 237-259.
- Piaget, J. (1932) The Moral Judgment of the Child. 1<sup>st</sup> ed. London. Kegan Paul, Trench, Trubner and Co. Ltd.
- Podlesny, J.A. and Raskin, D.C. (1977) Physiological Measures and the Detection of Deception. Psychological Bulletin (84) no.4, 782-799.
- Podlesny, J.A. (1993) Is the Guilty Knowledge Polygraph Technique Applicable in Criminal Investigations? A Review of FBI Case Records. Crime Laboratory Digest (20), 57-61.
- Porter, S. and Yuille, J.C. (1995) Credibility Assessment of Criminal Suspects Through Statement Analysis. Psychology, Crime and Law (1), 319-331.
- Porter, S. and Yuille, J.C. (1996) The Language of Deceit: An Investigation of the Verbal Clues to Deception in the Interrogation Context. Law and Human Behavior (20) no.4, 443-458.
- Posner, M.I. and Snyder, C.R.R (1975) Facilitation and Inhibition in the Processing of Signals. In P.M.A. Rabbit (ed.), Attention and Performance V. London, Academic Press.

- Postmes, T., Spears, R., Sakhel, K. and de Groot, D. (2001) Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior. Personality and Social Psychology Bulletin (27) no. 10, 1243-1254.
- Putnam, W.H. (1979) Hypnosis and Distortions in Eyewitness Memory. International Journal of Clinical and Experimental Hypnosis (27), 437-448.
- Quayle, E. and Taylor, M. (2001) Child Seduction and Self-Representation on the Internet. Cyberpsychology and Behavior (4) no.5, 597-608.
- Raskin, D.C. and Esplin, P.W. (1991a) Statement Validity Assessment: Interview Procedures and Content Analysis of Children's Statements of Sexual Abuse. Behavioral Assessment (13), 265-291.
- Raskin, D.C. and Esplin, P.W. (1991b) Assesment of Children's Statements of Sexual Abuse. In J. Doris (ed.) The Suggestibility of Children's Recollections. Washington D.C., American Psychological Association.
- Raskin, D.C. and Esplin, P.W. (1991c) Response to Wells, Loftus, and McGaugh. In J. Doris (ed.) The Suggestibility of Children's Recollections. Washington D.C., American Psychological Association.
- Raskin, D.C. and Podlesny, J.A. (1979) Truth and Deception: A Reply to Lykken. Psychological Bulletin (86) no.1,54-59.
- Ratcliff, R. and McKoon, G. (1981) Automatic and Strategic Priming in Recognition. Journal of Verbal Learning and Verbal Behavior (20), 204-215.
- Ratcliff, R. and Layton, W.M. (1981) A Microcomputer Interface for Control of Real-Time Experiments in Cognitive Psychology. Behavior Research Methods and Instrumentation (13), 216-220.
- Reicher, S.D. (1984) Social Influence in the Crowd: Attitudinal and Behavioural Effects of Deindividuation in Conditions of High and Low Group Salience. British Journal of Social Psychology (23), 341-350.
- Reicher, S.D. and Levine, M. (1994) Deindividuation, Power Relations between Groups and the Expression of Social Identity: The Effects of Visibility to the Out-Group. British Journal of Social Psychology (33) Part 2, 145-163.

Reid, J.E. and Inbau, F.E. (1977) Truth and Deception: The Polygraph ("Lie-Detector") Technique, 2<sup>nd</sup> ed. Baltimore, Williams and Wilkins.

Reid, J.E. (1947) A Revised Questioning Technique in Lie-Detection Tests. Journal of Criminal Law, Criminology and Police Science (37), 542-547.

Rheingold, H. (1993) The Virtual Community: Homesteading on the Electronic Frontier. MA, Addison-Wesley.

Riggio, R.E. and Friedman, H.S. (1986) Impression Formation: The Role of Expressive Behavior. Journal of Personality and Social Psychology (50) no.2, 421-427.

Rintel, E.S. and Pittam, J. (1997) Strangers in a Strange Land: Interaction Management on Internet Relay Chat. Human Communication Research (23) 507-534.

Rintel, E.S., Mulholland, J., and Pittam, J. (2001) First Things First: Internet Relay Chat Openings. Journal of Computer-Mediated Communication (6) no.2.

Roberts, L.D. and Parks, M.R. (1999) The Social Geography of Gender-Switching in Virtual Environments on the Internet. Information, Communication, and Society (2) no. 4, 521-540

Routtenberg, A. (1971) Stimulus Processing and Response Execution: A Neurobehavioral Theory. Physiology and Behavior (6), 589-596.

Ruby, C.L. and Brigham, J.C (1997) The Usefulness of the Criteria-based Content Analysis Technique in Distinguishing Between Truthful and Fabricated Allegations: A Critical Review. Psychology, Public Policy, and Law (3) no.4, 705-737.

Ruby, C.L. and Brigham, J.C. (1998) Can Criteria-Based Content Analysis Distinguish Between True and False Statements of African-American Speakers? Law and Human Behavior (22) no.4, 369-388.

Ryan, M.G. (1976) The Influence of Teleconferencing Medium and Status on Participants' Perception of the Aestheticism, Evaluation, Privacy, Potency, and Activity of the Medium. Human Communication Research (2) no.3, 255-261.

Sacks, H., Schegloff, E.A., and Jefferson, G. (1974) A Simplest Systematics for the Organisation of Turn-Taking for Conversation. Language (50), 696-735.

Sapir, A. (1987) The LSI Course on Scientific Content Analysis (SCAN). Phoenix, Laboratory for Scientific Interrogation. *Citation provided in (Shuy, 1998).*

Schiano, D.J. (1997) Convergent methodologies in Cyber-psychology: A Case Study. Behavior Research Methods, Instruments, and Computers (29) no.2, 270-273.

Schmitt, B.H. (1987) The Ecological Approach to Social Perception: A Conceptual Critique. Journal for the Theory of Social Behaviour (17) September. 265-278.

Seymour, T.L., Seifert, C.M., Shafto, M.G., and Mosmann, A.L. (2000) Using Response Time Measures to Assess "Guilty Knowledge". Journal of Applied Psychology (85) no.1, 30-37

Sheier, M.F. and Carver, C.S. (1977) Self-Focused Attention and the Experience of Emotion: Attraction, Repulsion, Elation, and Depression. Journal of Personality and Social Psychology (35), 625-636.

Short, J., Williams, E., and Christie, B. (1976) The Social Psychology of Telecommunications. 1<sup>st</sup> ed. Bath, Pittman Press.

Shuy, R.W. (1998) The Language of Confession, Interrogation, and Deception. 1<sup>st</sup> ed. Thousand Oaks, Sage.

Siddle, D.A.T. and Packer, J.S. (1991) Memory and Autonomic Activity: The Role of the Orienting Response. In J.R. Jennings and M.G.H.Coles (eds.) Handbook of Cognitive Psychophysiology: Central and Autonomic Nervous System Approaches. 1<sup>st</sup> ed. Chichester, John Wiley and Sons.

Siddle, D.A.T, Stephenson, D. and Spinks, J.A. (1983) Elicitation and Habituation of the Orienting Response. In. D. Siddle (ed.) Orienting and Habituation: Perspectives in Human Research. 1<sup>st</sup> ed. Chichester, John Wiley and Sons.

Spears, R. and Lea, M. (1992) Social Influence and the Influence of the 'Social' in Computer-Mediated Communication. In. M. Lea (ed.) Contexts of Computer-Mediated Communication.

Spears, R., Lea, M., and Lee, S. (1990) De-individuation and Group Polarization in Computer-Mediated Communication. British Journal of Social Psychology (29), 121-134

Spinks, J.A. and Siddle, D. (1983) The Functional Significance of the Orienting Response. In D. Siddle (ed.) Orienting and Habituation: Perspectives in Human Research. Chichester, John Wiley & Sons.

Spitzer, M. (1986) Writing Style in Computer Conferences. IEEE Transactions on Professional Communications, (29) no.1, 19-22.

Sporer, S.L. (1997) The Less Travelled Road to Truth: Verbal Cues in Deception Detection in Accounts of Fabricated and Self-Experienced Events. Applied Cognitive Psychology (11), 373-397.

Sporer, S.L. and Kuepper, B. (1995) Realitätsüberwachung und die Beurteilung des Wahrheitsgehaltes von Erzählungen: Eine experimentelle Studie [Reality Monitoring and the Judgement of Credibility of Stories: An Experimental Investigation]. Zeitschrift für Sozialpsychologie, (26), 173-193. *Citation provided in Sporer, S.L. (1997)*

Sporer, S.L., Kuepper, B. and Bursch, S.E. (1995) Hilft Wissen über Realitätsüberwachung, um Zwischen Wahren und Erfundenen Geschichten zu Unterscheiden? [Does Knowledge About Reality Monitoring Help to Discriminate Between True and Invented Stories?] Paper Presented at the 37<sup>th</sup> Tagung Experimentell Arbeitender Psychologen in Bochum. *Citation provided in Sporer, S.L. (1997)*

Sproull, L. and Kiesler, S. (1986) Reducing Social Context Cues: Electronic Mail in Organisational Communication. Management Science (32) no. 11, 1492-1512.

Sproull, L. and Kiesler, S. (1991) Connections: New Ways of Working in the Networked Organisation. Cambridge, MIT Press.

Steinfeld, C.W. (1986) Computer-Mediated Communication in an Organisational Setting: Explaining Task-Related and Socioemotional Uses. In M.L. McLaughlin (ed.) Communication Yearbook 9. Newbury Park, Sage.

Steller, M. (1989) Recent Developments in Statement Analysis. In J.C. Yuille (ed.), Credibility Assessment, Kluwer Academic Publishers.

Steller, M. and Köhnken, G. (1989) Criteria-Based Content Analysis. In D.C. Raskin (ed.) Psychological Methods in Criminal Investigation and Evidence. 1<sup>st</sup> ed. New York, Springer-Verlag.



Steller, M. and Wellershaus, P. (1996) Information Enhancement and Credibility Assessment of Child Statements: The Impact of the Cognitive Interview on Criteria-Based Content Analysis. In G. Davis, S. Lloyd-Bostock, M. McMurrin, and C. Wilson (eds) Psychology, Law, and Criminal Justice: International Developments in Research and Practice. Berlin, de Gruyter.

Stern, W. (1902) Zur Psychologie der Aussage. [To the Psychology of Testimony]. Zeitschrift für die Gesamte Strafrechtswissenschaft, (22). *Citation provided in (Undeusch, 1989)*.

Stiff, J.B. (1995) Conceptualizing Deception as a Persuasive Activity. In C.R. Berger and M. Burgoon (eds.) Communication and Social Influence Processes. East Lansing, Michigan State University Press.

Stiff, J.B. (1996) Theoretical Approaches to the Study of Deceptive Communication: Comments on Interpersonal Deception Theory. Communication Theory (6) no.3, 289-295.

Svenning, L.L. and Ruchinskas, J.E. (1984) Organizational Teleconferencing. In R.E. Rice and Associates (eds), The New Media: Communication, Research, and Technology. Beverly Hills, Sage.

Sweetland Films (1998) The Spanish Prisoner. Produced by Jean Doumanian. Directed by David Mamet. © Sweetland Films.

Tarbox, K. (2001) Katie.com: One Girl's Loss of Innocence. Orion.

Thompson, P.A. and Foulgar, D.A. (1996) Effects of Pictographs and Quoting on Flaming in Electronic Mail. Computers in Human Behavior (12) no. 2, 225-243

Tully, B. (1998) Reliability of Criteria-based Content Analysis of Child Witness Statements: Cohen's Kappa Doesn't Matter. Legal and Criminological Psychology (3), 183-188.

Tulving, E. and Craik, F.I.M. (2000) The Oxford Handbook of Memory. 1<sup>st</sup> ed. Oxford, Oxford University Press.

Turkle, S. (1995) Life on the Screen: Identity in the Age of the Internet. 1<sup>st</sup> ed. New York, Simon & Schuster.

Turkle, S. (1996) Parallel Lives: Working on Identity in Virtual Space. In D. Grodin and T.R. Lindlof (eds.) Constructing the Social Self in a Mediated World. New York, Sage.

Turnbull, W. (2003) Language in Action: Psychological Models of Conversation. 1<sup>st</sup> ed. Hove, Psychology Press.

Turner, J.C. (1987) A Self-Categorisation Theory. In. J.C. Turner, M.A. Hogg, P.J. Oakes, S.D. Reicher, and M.S. Wetherell Rediscovering the Social Group: A Self-Categorisation Theory. 1<sup>st</sup> ed. Oxford, Blackwell.

Undeutsch, U. (1989) The Development of Statement Reality Analysis. In. J.C. Yuille (ed.), Credibility Assessment, Kluwer Academic Publishers.

Utz, S. (2000) Social Information Processing in MUDs: The Development of Friendships in Virtual Worlds. Journal of Online Behavior (1), no.1.

Van Gelder, L. (1996) The Strange Case of the Electronic Lover. In. R. Kling (ed.) Computerization and Controversy: Value Conflicts and Social Choices. 2<sup>nd</sup> ed. London, Academic Press.

Vasek, M.E. (1986) Lying as a Skill: The Development of Deception in Children. In R.W. Mitchell and N.S. Thompson (eds.) Deception: Perspectives on Human and Nonhuman Deceit. 1<sup>st</sup> ed. Albany, State University of New York Press.

Vrij, A., Edward, K., Roberts, K.P., and Bull, R. (2000) Detecting Deceit via Analysis of Verbal and Nonverbal Behavior. Journal of Nonverbal Behavior (24) no.2. Winter, 239-263.

Vrij, A. (2000) Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice. 1<sup>st</sup> ed. Chichester, John Wiley & Sons Ltd.

Walther, J.B. and Tidwell, L.C. (1995) Nonverbal Cues in Computer-Mediated Communication, and the Effect of Chronemics on Relational Communication. Journal of Organisational Computing (5) no. 4, 355-378

Walther, J.B. (1992) Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. Communication Research (19) no.1. February, 52-90

Walther, J.B. (1993) Impression Development in Computer-Mediated Interaction. Western Journal of Communication (57) Fall, 381-398

- Walther, J.B. (1994) Anticipated Ongoing Interaction Versus Channel Effects on Relational Communication in Computer-Mediated Interaction. Human Communication Research (40), 473-501.
- Walther, J.B. (1996) Computer Mediated Communication: Impersonal, Interpersonal and Hyperpersonal Interaction. Communication Research (23) no. 1. February, 3-43
- Waskull, D. and Douglass, M. (1997) Cyberself: The Emergence of Self in On-Line Chat. The Information Society (13), 375-397
- Waterfield, R. (1998) Republic. Plato, translated by R. Waterfield. Oxford, Oxford University Press.
- Weston, J.R. and Kristen, C. (1973) Teleconferencing: A Comparison of Attitudes, Uncertainty and Interpersonal Atmospheres in Mediated and Face-to-Face Group Interaction. The Social Policy and Programs Branch, Department of Communications, Ottawa, Ontario, Canada, Report No. 3, Contract OSU4-0072. *Citation provided in (Johansen et al., 1979).*
- White, C.H. and Burgoon, J.K. (2001) Adaptation and Communicative Design: Patterns of Interaction in Truthful and Deceptive Conversations. Human Communication Research (27) no.1, 9-37.
- Whitty, M. and Gavin, J. (2001) Age/Sex/Location: Uncovering the Social Cues in the Development of Online Relationships. Cyberpsychology and Behavior (4) no. 5, 623-694
- Whitty, M.T. (2002) Liar, Liar! An Examination of How Open, Supportive and Honest People are in Chat Rooms. Computers in Human Behavior (18), 343-352
- Wilbur, S.P. (1997) An Archaeology of Cyberspaces: Virtuality, Community, Identity. In D. Porter (ed.) Internet Culture. New York, Routledge.
- Wiener, M. and Mehrabian, A. (1968) Language Within Language: Immediacy, a Channel in Verbal Communication. New York, Appleton-Century-Crofts.
- Zahavi, A. (1993) The Fallacy of Conventional Signalling. The Royal Society Philosophical Transaction B (340), 227-230.
- Zaparniuk, J., Yuille, J.C., and Taylor S. (1995) Assessing the Credibility of True and False Statements. International Journal of Law and Psychiatry (18) no.3, 343-352.

Zebrowitz, L.A. (1990) Social Perception. 1<sup>st</sup> ed. Milton Keynes, Open University Press.

Zhou, L., Twitchell, D.P., Qin, T., Burgoon, J.K., and Nunamaker, J.F. (2003) An Exploratory Study into Deception Detection in Text-based Computer-Mediated Communication. Paper presented at the 36<sup>th</sup> Hawaii International Conference on System Sciences.

Zimbardo, P.G. (1969) The Human Choice: Individuation Reason and Order Versus De-individuation Impulse and Chaos. In W.J. Arnold and D. Levine (eds) Nebraska Symposium on Motivation (17) Lincoln, University of Nebraska Press.

Zimmerman, D.P. (1987) Effects of Computer Conferencing on the Language Use of Emotionally Disturbed Adolescents. Behavior Research Methods, Instruments, and Computers (19) no. 2, 224-230

Zuckerman, M., DePaulo, B.M. and Rosenthal, R. (1981) Verbal and Nonverbal Communication of Deception. In L. Berkowitz (ed.) Advances in Experimental Social Psychology (14) 1-59.

